Characterizing the Relationship Between Cell-Cycle Progression and a Transcriptional Oscillator

by

Sara Lynn Bristow

University Program in Genetics and Genomics
Duke University

Date:_____

Approved:

_____
Steven Haase, Supervisor

_____
Blanche Capel

_____
Beth Sullivan

_____
David MacAlpine

_____
Thomas Petes

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy in the
University Program in Genetics and Genomics
in the Graduate School of Duke University

2013

ABSTRACT

Characterizing the Relationship between Cell-Cycle Progression and a Transcriptional Oscillator

by

Sara Lynn Bristow

University Program in Genetics and Genomics
Duke University

Date:_____

Approved:

_____
Steven Haase, Supervisor

_____
Blanche Capel

_____
Beth Sullivan

_____
David MacAlpine

_____
Thomas Petes

An abstract of a dissertation submitted in partial fulfillment
of the requirements for the degree of Doctor of Philosophy in the
University Program in Genetics and Genomics
in the Graduate School of Duke University

2013

Abstract

The cell division cycle is the process in which the entirety of a cell's contents is duplicated completely and then equally segregated into two identical daughter cells. The order of the steps in the cell cycle must be followed with fidelity to guarantee two viable cells. Understanding the regulatory mechanisms that control cell-cycle events remains to be a fundamental question in cell biology. In this dissertation, I explore the mechanisms that coordinate and regulate cell-cycle progression in the budding yeast, *Saccharomyces cerevisiae*.

Cell-cycle events have been shown to be triggered by oscillations in the activity of cyclin dependent kinases (CDKs) when bound to cyclins. However, several studies have shown that some cell-cycle events, such as periodic transcription, can continue in the absence of CDK activity. How are periodic transcription and other cell-cycle events coupled to each other during a wild-type cell cycle? Currently, two models of cell-cycle regulation have been proposed. One model hypothesizes that oscillations in CDK activity controls the timing of cell-cycle events, including periodic transcription. The second model proposes that a transcription factor (TF) network oscillator controls the timing of cell-cycle events, via proper timing of gene expression, including cyclins. By measuring global gene expression dynamics in cells with persistent CDK activity, I show that periodic transcription continues. This result fits with the second model of cell-cycle regulation. Further, I show that during a wild-type cell cycle, checkpoints

are responsible for arresting the bulk of periodic transcription. This finding adds a new layer of regulation to the second model, providing a mechanism that coordinates cell-cycle events with a TF network oscillator. Taken together, these data provide further insight into the regulation of the cell cycle.

Contents

List of Tables

List of Figures

Acknowledgements

I first have to thank my advisor, Steve, for all of his support and advice throughout my entire graduate career. Steve, you have challenged and pushed me far beyond what I thought I was capable of doing. You have taught me many valuable lessons, not only in science, but also in my professional and personal life. I value all of the discussions, jokes, and time spent with you. You have made the lab such a fun place to be, allowing me to not only work hard but enjoy my time while doing it. I am so thankful I never received that 'memo' all those years ago. I cannot imagine having a better advisor, mentor, and friend. I hope that I am able to inspire others to be their best, just as you have done for me.

I would also like to acknowledge my committee members, Beth, Blanche, Dave, and Tom. Thank you for always challenging me to continue to question my results. For help with a number of experiments, I thank Sam Johnson, Yasheng Gao, Holly Dressman, Laura-Leigh Rowlette, Sayan Mukherjee, and many others. I would also like to thank Carol Richardson, Andrea Lanahan, Leslie Mavengere, Anne Lacey, and all other administrative staff in the Biology Department for all of the help that they provided me.

I would like to thank everyone in the Haase lab, and my collaborators. Laura, you have been an incredible mentor and friend to me. I appreciate all of the advice and help you gave me, even after you moved on to UNC. Dave, thank you for making the lab such a fun place to be and being my first teacher in the art

of computational biology. Mark, thank you for being such a patient and excellent teacher and resource in the lab. Michael, thank you for not only teaching me about computational biology, but also about beer making. Chun-Yi, thank you for always knowing the right questions to push my work farther. Adam and Tina, I have had so much fun learning with you. I will always treasure our shenanigans and remember them with a smile on my face. Anastasia, thank you for being a great friend and collaborator. I will miss our knitting sessions together. To everyone else, thank you for all that you have done to help me.

I have been most fortunate to have made lifelong friends when I first arrived at Duke. We have shared both highs and lows, making the successes much more enjoyable, and the defeats a little more bearable. I have enjoyed our lunches, trips, and nights out together. Thank you for the memories.

And last, but for sure not the least, I need to thank my family. Dad, Melissa, Elliot, Kelsey, and Matt, thank you for all of your support. I do not think I could have done this without you. Andrew, thank you for walking through this journey with me from the beginning. Even though we lived apart from each other, you reveled in my successes whole-heartedly. I cannot wait to take my next steps in our journey together with you in New York.

Chapter 1 Introduction

Cellular division cycles serve as a fundamental biological process that underlies reproduction, development, evolution, and cancer. A complex number of steps must be completed with fidelity and in the correct order to ensure two exact copies of the same cell. Multiple layers of regulation play a role in successful cellular divisions. In this dissertation, I will explore the mechanisms that serve the basis of cell-cycle regulation and present evidence of how these regulatory modules are coupled to each other.

1.1 The cell cycle is a series of temporally ordered events

Successful cellular division requires complete duplication of genetic material followed by their equal segregation into two cell bodies, resulting in two identical daughter cells. Here, order is critical: duplication must be completed before division events commence. By observing cells during this process, the cell cycle has been divided into four phases - Gap 1 (G1), Synthesis (S), Gap 2 (G2), and Mitosis (M) - that describe the chronological order of different events observed in normal cycling cells. DNA duplication occurs during S phase. In addition to replicating chromosomes, the centrosome is also duplicated and the daughter centrosomes separated in order to form the mitotic spindle that is responsible for separating sister chromatids into two separate nuclei.

Subsequent sister chromatid segregation is initiated during M phase, or mitosis. G1 and G2 are referred to as ‘gap’ phases as they separate the visibly observable events of S- and M- phase.

Although no overt changes or events are occurring during G1 and G2, cells are actively interacting with their extracellular and intracellular environments to ensure that conditions are appropriate for cellular division events. During G1, cells are not committed to undergoing cellular division. Rather, at this point in a cell’s life, a choice between fates is possible and is dependent on the extracellular environment. When nutrients and space are in abundance, cells initiate entry into the cell cycle. However, when nutrients and space are in short supply, cells do not divide and may enter quiescent phase, also referred to as G0 phase. Alternatively, budding yeast cells can also initiate other cell-fate programs, such as pseudohyphal growth and meiosis. The quiescent phase is reversible; cells can re-enter G1 phase and commit to a cellular division when conditions change. When a favorable environment for cellular division is sensed, cells commit to the cell cycle by transitioning from G1 phase to S phase. This transition is referred to as Start in the budding yeast, *Saccharomyces cerevisiae*, and the Restriction Point in metazoan cells. G2 serves as a time to ensure that DNA replication has been completed with fidelity and that the spindle is correctly formed.

1.2 Identifying the biochemical regulators of cell-cycle progression

The order of events is very well conserved in unperturbed cells performing multiple rounds of cellular divisions. Each cell-cycle event – such as DNA replication, centrosome duplication, and chromosome segregation – is a complex process that requires the coordination of many different proteins acting together to complete the task at hand. These complicated mechanisms must be initiated precisely in the context of the cell cycle. How is it that DNA replication always initiates only once per cycle and always before mitosis? Mammalian cell fusion studies have shown that S-phase nuclei are able to trigger DNA replication in G1 nuclei, suggesting the presence of some factor that activates DNA replication [1]. However, inhibitory factors also exist that prevent late events from occurring prior to early events. G2 nuclei fused with either S phase nuclei or G1 nuclei wait to undergo mitosis until DNA replication is completed in the S-phase or G1 nuclei [1]. What are these cellular components and how do they function to activate or inhibit cell-cycle events?

1.3 Cyclins and Cyclin dependent kinases

A series of studies in early developing marine embryos began to elucidate the factors involved in regulating cell-cycle events. Maturation promoting factor (MPF) was determined to be responsible for driving rapid duplication and division cycles in early embryos; MPF was subsequently determined to be composed of

cyclin and cyclin-dependent kinase (CDK) [2, 3]. Cyclins bind to and activate CDKs that modulate the activity of proteins involved in the cell cycle by phosphorylation (reviewed in [4, 5]). In early developing embryos, cyclin protein synthesis is periodic with respect to the cell cycle, progressively increasing during S phase and DNA replication until cyclin protein is abruptly destroyed just prior to mitosis and chromosome segregation [6]. What causes periodic cyclin protein levels in these embryos? Cyclin/CDK activates its own inhibitor, an E3 ubiquitin ligase called the anaphase-promoting complex (APC) [7, 8]. This biochemical negative feedback loop between cyclin/CDK and the APC constitutes the cell-cycle oscillator in early developing marine embryos.

In somatic metazoan cells, oscillations in cyclin/CDK activity are also found to initiate different cell-cycle events. The same APC-mediated negative feedback loop identified in embryos is thought to act as the cell-cycle oscillator in these somatic cells as well. However, this model becomes more complicated as multiple cyclins and CDKs act at different times during the somatic cell cycle [9-11]. The cyclin originally identified in embryos is referred to as cyclin B, and the original CDK identified is now called Cdk1 and is responsible for activating mitosis. Cyclin D complexed with Cdk4/6 controls cell cycle entry at the Restriction Point while cyclins A and E activate Cdk2 to initiate DNA replication in somatic metazoan cells (reviewed in [4, 5]).

Concurrent with the above-described studies, genetic approaches to elucidating the mechanism of cell-cycle control were carried out in the budding yeast, *Saccharomyces cerevisiae* [12-16], and the fission yeast, *Schizosaccharomyces pombe* [17]. Genes homologous to cyclins and CDK in budding and fission yeasts were identified. A single CDK was identified in each of these microorganisms, Cdc28 in budding yeast and cdc2 in fission yeast. Studies show that both possess kinase activity [18, 19], are homologous to each other [20, 21], and complement each other [22]. Further, human *CDC2* cDNA is able to a complement *cdc2* fission yeast deletion mutant, suggesting that CDK is strongly conserved across species [23]. A total of nine cyclins have been identified to play a role in activating Cdc28 in budding yeast (Figure 1.1).

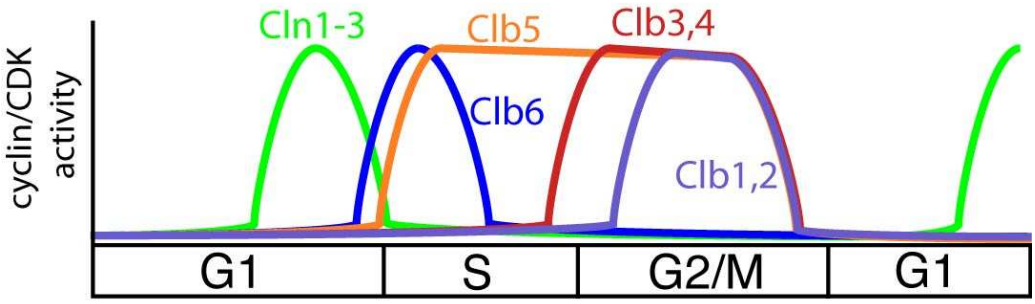

Figure 1.1: Timing of cyclin/CDK activity in Saccharomyces cerevisiae. Different cyclins are expressed throughout the cell cycle to trigger different events. Each set of cyclins is shown in different colors.

Three G1 cyclins, Cln1, Cln2 and Cln3, regulate passage through Start and G1 entry, thus committing the cell to a new division cycle. The six cyclins that are homologous to metazoan cyclin B, Clb1, Clb2, Clb3, Clb4, Clb5 and Clb6

*(*referred to as B-type cyclins), regulate the cell-cycle events initiated after Start. Clb5 and Clb6 trigger DNA replication, while Clb1-4 play roles in chromosome segregation (reviewed in [4, 5]). Additionally, the APC has also been identified in budding yeast [24, 25]. The conservation of the components and function, tt has been proposed that the same biochemical oscillator described in marine embryos also regulates cell-cycle progression in yeast cells (reviewed in [26]).

1.4 Saccharomyces cerevisiae as a model system for studying the cell cycle

As such, this high level of conservation of cell-cycle regulatory components makes the budding yeast, *Saccharomyces cerevisiae* an ideal model system to study the cell cycle. Additionally, cell-cycle phases in budding yeast can be inferred by changes in cell morphology that are easily observed by light microscopy (Figure 1.2).

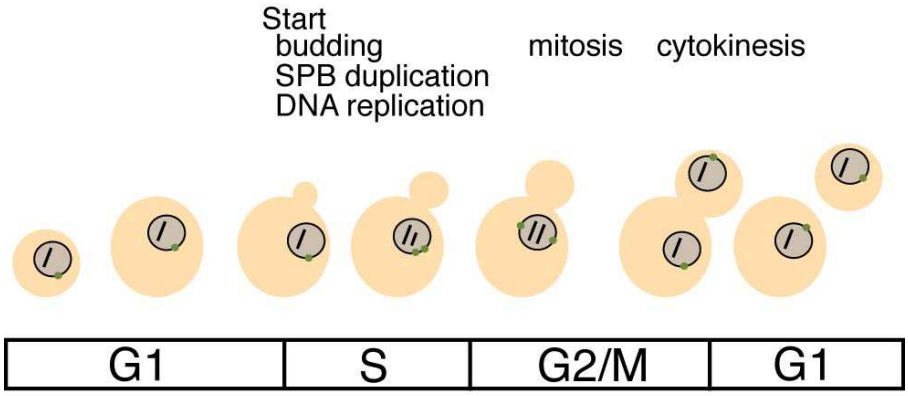

Figure 1.2: Budding yeast as a model system for studying the cell cycle. Top: The timing of cell-cycle events. Bottom: The phenotype of budding yeast as it progresses through the cell cycle. Black lines, DNA; Green dots, spindle pole bodies (SPBs)

Just prior to Start, the budding yeast cell produces a projection called a bud that is destined to become the daughter cell, where approximately half of the cellular contents are transferred during mitosis. This bud continues to grow as the cell progresses through the cell cycle. Thus, unbudded cells are in G1, and cells that have formed a bud have entered the cell cycle, where the size of the bud is an indicator of cell-cycle phase; small budded cells are likely to be in S-phase while large budded cells are undergoing mitosis (Figure 1.2).

Budding yeast has been used as a model system to understand how cyclin/CDK activity during the cell cycle is responsible for triggering DNA replication, centrosome duplication (referred to as spindle pole bodies in budding yeast), and mitosis, in addition to many other events. Cyclin/CDKs phosphorylate proteins to modulate their activity, both positively in order to promote their activity, and negatively to inhibit their activity. For instance, S-phase cyclins activate DNA replication, but also inhibit the re-initiation of DNA replication until the following cell cycle begins [27]. Cyclins function in a similar manner to prevent multiple rounds of centrosome duplication [28, 29].

1.5 Checkpoints act as a control module to couple cell-cycle events to each other

Perhaps the most critical aspect of cell-cycle control is ensuring that different events occur in the proper temporal order. Different cyclin/CDK activities are responsible for triggering events throughout the cell cycle. What happens

when an event early in the cycle cannot be completed in time before later events are normally initiated? Some mechanism must be responsible for ensuring that late cell-cycle events are not triggered until early cell-cycle events are completed with fidelity. Checkpoints are signaling pathways that monitor the progression through different events (reviewed in [30]. Checkpoint signaling pathways utilize three main classes of proteins: sensors, transducers, and effectors. Once a perturbation is identified, a sensor protein is activated and goes on to trigger transducer proteins that carry the signal to effector proteins. The effector proteins are responsible for inducing the cell-cycle arrest. Once the perturbation is relieved, the sensor protein is no longer active, resulting in the transducer and effector proteins to no longer affect cell-cycle progression.

A different checkpoint-signaling pathway monitors each cell-cycle event to maintain the proper order of events. The morphogenesis checkpoint monitors bud emergence, and inhibits mitosis until the bud can be formed (reviewed in [31]). The DNA damage checkpoint can arrest cell-cycle progression any time single- or double-stranded breaks are sensed (reviewed in [32, 33]). The DNA replication checkpoint ensures that DNA is fully replicated during S-phase prior to initiating mitosis (reviewed in [32, 33]. Finally, the spindle assembly checkpoint prevents chromosome segregation if the orientation of the microtubules extended from the mitotic spindle are not attached to the sister chromatids properly (reviewed in [34]). In this way, two layers of regulation have been shown to play a

role in ensuring that during each cell cycle, all events are completed with fidelity and in the proper order.

1.5.1 The DNA replication checkpoint signaling pathway

The DNA replication checkpoint is thought to sense stalled replication forks during the replication process [35] (reviewed in [36]). In fact, many of the components of the replication machinery also function to signal when DNA replication is stalled [37-40]. Once these proteins are activated, they pass this signal to the sensor protein Mec1, a kinase, that signals to downstream transducers and effectors [41, 42]. Mec1 phosphorylates and activates another kinase, Rad53, which functions as both a transducer and effector [43]. Rad53 plays many roles during the DNA replication checkpoint.

First, and most importantly, Rad53 phosphorylates and inhibits Cdc20, the activating subunit of the APC [44]. By inhibiting APC^{Cdc20} activity, many proteins are stabilized for the duration that the DNA replication checkpoint is active. Stabilization of APC target Pds1, also known as securin, prevents chromosome segregation by inhibiting the protein separase, Esp1 [45, 46]. In addition to Pds1, mitotic cyclin Clb2 is also a target of the APC and is stabilized during the checkpoint [47, 48]. As a result, Clb2/CDK activity is persistent throughout the DNA replication checkpoint [47, 48]. Previous studies have shown that overexpression of a hyper-stable Clb2 in normally cycling cells arrests cell-cycle

progression, indicating another mechanism that may aid in the checkpoint-induced arrest [49].

A second role for Rad53 is to aid in overcoming the replicative stress faced by the cells. Often the DNA replication is induced when nucleotide pools are depleted. Rad53 works to help increase the availability of free nucleotides by activating the checkpoint kinase and effector Dun1 [50]. Dun1 kinase functions to activate the enzyme ribonucleotide reductase, the rate-limiting step in producing new pools of dNTPs. Dun1 acts to increase pools of dNTPs in two ways. First, Dun1 activates ribonucleotide reductase by inhibiting its inhibitor, Sml1 [51, 52]. Second, Dun1 phosphorylates and inactivates the transcriptional repressor Rfx1 that is responsible for repressing *RNR3* gene expression [53-56]. In addition to *RNR3*, Dun1 has been shown to induce the expression of a number of other genes (reviewed in [57-59]. Thus, the DNA replication checkpoint arrests cell-cycle progression, aids in overcoming replicative stress, and induces a checkpoint-specific transcriptional response.

1.5.2 The spindle assembly checkpoint signaling pathway

The spindle assembly checkpoint monitors spindle attachment to sister chromatids and ensures a bipolar orientation so that the resulting daughter cells receive one copy of each chromosome. Similarly to the DNA replication checkpoint, the mechanism of preventing sister chromatid segregation is based on inactivating Cdc20 (reviewed in [34]). This inhibition is not achieved by

phosphorylation, but rather by physical interactions between Cdc20 and checkpoint proteins. There are a number of proteins that have been shown to physically interact with Cdc20, and it is as yet unclear the exact order of signaling and binding that leads to APC inactivation [60-63]. These checkpoint proteins include Mps1, Mad1, Mad2, Mad3, Bub1, and Bub3. These proteins are concentrated at kinetochores, where they are able to trigger the spindle assembly checkpoint.

Unlike, the DNA replication checkpoint, no checkpoint-specific transcriptional response has been found to be activated along with the spindle assembly checkpoint. Thus, it appears that the main action of the spindle assembly checkpoint is to delay the onset of mitosis until proper spindle-kinetochore attachment is achieved.

1.6 Cell-cycle regulated transcription

The cyclin/CDK complex was the first cell-cycle regulator shown to have oscillatory activity that can direct event timing, but it is not the only mechanism of cell-cycle control. Another phenomenon that has been observed during the cell cycle in both budding yeast and mammalian cells is periodic transcription, in which genes are expressed during only one phase of the cell cycle. How many genes are regulated at the transcriptional level during the cell cycle? As all other cell-cycle events are controlled by cyclin/CDK activity, is periodic transcription also regulated similarly?

1.6.1 Identifying periodic transcripts using Northern blots

With the advent of modern molecular biology, measuring mRNA levels in cells became a regular test to address whether genes are regulated at the transcriptional level. For genes involved in the cell cycle, understanding gene regulation at the transcriptional level requires measuring mRNA in populations of cells that have been synchronized to one phase of the cell cycle and subsequently released to determine how gene transcript levels changes or does not change. Histones were the first genes identified that are expressed periodically during the cell cycle [64]. Classifying histone gene expression as cell-cycle regulated was done by correlating the timing of histone mRNA expression with the timing of DNA replication over the course of several cell cycles [64]. Over the following decade, ten more genes involved in cell-cycle events were also identified as being expressed in a periodic manner – *HO* [65], *CDC21* [66], *CDC9* [67], *RAD6* [68], *SWI5* [69], *CDC8* [70], *POL1* [71], *DBF4* [72], *PRI1* [73], and *DBF2* [74]. For each of these periodic genes, the definition of periodic is anchored to the correlation of gene expression with an observable cell-cycle event that is known to occur only once per cycle. The periodic expression of these genes was discovered while investigating the function of each gene during cell-cycle events. Is periodic expression of cell-cycle genes a global phenomenon or specific to just a small set of genes? The approach to address this question is

to survey transcript dynamics on a large scale in synchronous populations of cells.

Northern blots are best used to measure mRNA expression dynamics on an individual gene basis. However, attempting to determine the extent to which genes are periodically expressed during the cell cycle requires the ability to screen through all genes in the genome. To address the question posed above using northern blots, Price and colleagues designed a screen-based approach to identify cell-cycle regulated transcripts in a high-throughput manner [75]. In this study, genes were first screened for potential periodic transcriptional dynamics in populations of cells synchronized to different cell-cycle phases. Transcripts found to be expressed at only one phase of the cell cycle were then measured in wild-type synchronous populations of cells over time. This method identified at least 20 additional periodic genes and further estimated that up to 250 of the 6,000 budding yeast genes are cell-cycle regulated based on these findings [75]. This study revealed a large number of potential transcriptionally-regulated genes during the cell cycle. However, due to the limitations of northern blotting, discussed further in a later section, it is evident that more than 250 genes are periodically expressed.

Outside of this screen-based approach, further single gene studies identified additional periodic genes that are expressed concurrently with cell-cycle events. In total, approximately 100 periodically expressed budding yeast

genes were identified. However, a technology that is able to measure gene expression dynamics at a genome-wide level was necessary not only to quantify the proportion of genes that is periodically transcribed but also to understand how cell-cycle regulated transcription is coordinated with cell-cycle progression.

1.6.2 Identifying periodic transcripts using gene expression microarrays

The gene expression microarray is one method developed to measure mRNA levels of many genes in an organism [76]. This experimental approach serves as a useful tool for identifying cell-cycle regulated genes. Several genome-wide studies have been reported that focused on identifying periodic genes with respect to the cell cycle in synchronized populations of budding yeast cells using microarrays. This approach requires multiple microarrays to measure global mRNA levels at multiple time points in the synchronous population of cells. Cho and colleagues identified 416 genes as being cell-cycle regulated at the transcriptional level by visual inspection of transcript abundance over time [77]. This study used two different temperature-sensitive mutants to generate synchronous populations of cells at different phases of the cell cycle – the *cdc28-13* allele arrests cells in G1 while the *cdc15-2* allele arrests cells in late M phase. Spellman and colleagues identified 800 genes that demonstrate oscillations in gene expression during the cell cycle using quantitative methods including a Fourier transform and Pearson correlation [78]. These two quantifications are

then combined into an aggregate CDC score that ranks the genes based on periodicity [78]. To further demonstrate that the expression dynamics are a general phenotype of cycling cells, this group used three synchronization methods to measure global gene expression dynamics and to identify periodic transcripts: alpha factor to synchronize cells in late G1, *cdc15-2* to synchronize cells in late M phase, and centrifugal elutriation to synchronize cells in early G1.

Pramila and colleagues found 991 cell-cycle regulated transcripts using a permutation-based method developed by Ulrik de Lichtenberg and colleagues [79, 80]. This approach quantifies periodicity using a Fourier transform and amplitude using peak-to-trough ratios. These two scores are combined into statistical p-values with random permutations of the datasets and then integrated to rank the genes by periodicity. This study synchronized cells in late G1 with alpha factor and performed two biological replicates. Orlando and colleagues identified 1275 periodically expressed genes also using a permutation-based method [79, 81]. This study isolated small, unbudded cells in early G1 using centrifugal elutriation and also measured gene expression dynamics in two biological replicates. Between the three studies using quantitative methods to identify periodic genes, 440 cell-cycle regulated genes are shared [78, 80, 81]. Even though each study identifies slightly different sets of periodic genes, it is clear that many more genes are regulated at the transcriptional level during the cell cycle than previously thought. Differences between identified periodic gene

lists from each study result from a combination of experimental design and quantitative analysis. Further discussion of these differences is addressed in a later section.

Is the phenomenon of cell-cycle regulated transcription specific only to budding yeast? Additional studies in fission yeast and human cells have measured gene expression dynamics in synchronized cells to determine the scope of periodic transcription in these organisms. In fission yeast, three genome-wide studies identified a limited number of periodic genes [82-84]. Unlike budding yeast, fewer genes were classified as cell-cycle regulated; less than 800 genes were identified by each study. This result may be due to the synchronization methods used by each group. Similar to budding yeast, the agreement of each study is very low, with only 171 genes shared between all analyses [82-84]. Two studies in human cell lines have classified fewer than 1,000 periodic genes [85, 86]. The low numbers of periodic genes may be due to the larger genome size, undetected alternative splicing of introns, or the difficulty involved in synchronizing fission yeast or human cell lines. While budding yeast has the largest number of identified periodic genes, cell-cycle regulated transcription is also clearly observed in fission yeast and human cell lines, suggesting that this phenomenon is conserved between organisms. With improved experimental approach and mRNA measuring technology, the

characterization of periodic gene expression will become more tractable in other model systems.

Two questions arise from these striking findings: (1) what is the significance of cell-cycle regulated transcription and (2) what mechanisms coordinate this large transcriptional program with cell-cycle progression?

1.7 Significance of cell-cycle regulated transcription

Many hypotheses have been posed to explain the importance of cell-cycle regulated transcription. All of these explanations can be generalized into four categories. While all are plausible reasons for regulating gene expression timing during the cell cycle, it is not currently possible to discriminate between the multiple hypotheses. Moreover, each potential hypothesis is not mutually exclusive and may be true for only a subset of cell-cycle regulated transcription.

The first explanation postulates that cell-cycle regulated transcription is a mechanism to expend energy resources efficiently, as transcription and translation are energetically expensive. This concept is often referred to as “just in time” transcription, in which gene products that function at a specific cell-cycle interval are expressed only when needed (reviewed in [87-89]). A variation on this first explanation has been referred to as the “Sleeping Beauty” situation, which takes into account the full lifetime of a cell or tissue, rather than the completion of a single cell cycle ([78], reviewed in [90]). Whether it is a single budding yeast or a population of cells that form tissues in an organism, active

cellular division occurs during only a portion of a cell's overall life cycle. Microorganisms are subject to environmental constraints and will divide when conditions permit it but not when the local environment (nutrients, temperature, growth factor signaling, etc) is not amenable to cell division. Therefore, much of the life of a single cell is spent outside the cell cycle, in a state of rest or quiescence. During embryotic development, cells divide multiple times to give rise to tissues that result in a multicellular organism. However, after tissue development, most cells no longer receive the signal to divide and instead carry out functions specific to each tissue type. While these cells are no longer dividing, there is no need to expend energy expressing genes required for cell-cycle progression. However, once a signal is received to initiate cellular division, the cells are poised to complete cell-cycle events with the proper genes expressed at the correct time.

A second explanation for the function of periodic transcription relates to proper timing of cell-cycle regulators involved in triggering subsequent events. Activators of gene expression need to be "ON" at discrete times, in the absence of repressors directing expression to be "OFF." Otherwise, expression would not occur at all. Examples include cyclins, inhibitors of cyclins, and other genes that trigger events to initiate at the proper time (reviewed in [87, 88]). This concept can be made clear by constitutively expressing a hyper-stable version of the S-phase and mitotic cyclin inhibitor, Sic1. If Sic1 is always expressed and inhibits

S-phase and mitotic cyclins, cells would be blocked from initiating DNA replication or mitosis, which results in the arresting of cell-cycle progression [91]. Sic1 must be expressed only after mitosis and during the subsequent G1 phase to prevent early activation of DNA replication or mitotic events [92, 93]. In addition to transcriptional control of Sic1, post-translational control also exists to inactivate and degrade Sic1 at the proper time [91, 94]. Thus, periodic transcription plays an important role in regulating the timing and coordination of cell-cycle events.

A third proposal for the importance of cell-cycle regulated transcription centers on building a required structure only once per cell cycle (reviewed in [87, 89]). For example, proteins required for DNA replication are loaded onto DNA in different stages. The components of the replication complex are periodically transcribed themselves, lending to the temporal events that are required for DNA replication. A pre-initiation complex first binds to DNA replication origins and is only activated when elements are phosphorylated by S-phase cyclin/CDK. Other components required for replication are then synthesized, recruited to origins, and replicate DNA (reviewed in [95]). Further, mitotic cyclin/CDK activity inhibits the reformation of pre-initiation complexes until the following cell cycle [27]. This mechanism ensures that the complex required to trigger DNA replication is only built once and thus acts as a layer of control to prevent re-replication.

A fourth reason for periodic transcription centers on renewing pools of unmodified protein. Gene products that are post-translationally modified may no longer be active or be responsive to additional signaling. Therefore, periodic transcription provides a pool of unmodified product that is able to carry out cell-cycle events (reviewed in [90]). For example, Swi6, a component of transcription factor complexes SBF and MBF, is phosphorylated in S phase after Start to localize it to the cytoplasm [96]. Periodic transcription of *SWI6* may provide a new pool of Swi6 protein to induce transcription at Start. Four varying, but valid hypotheses exist to explain the phenomenon of cell-cycle regulated transcription. The underlying requirement for proper expression timing during the cell cycle has led to the development of a complicated program for cell-cycle transcription control. Understanding how periodic transcription is regulated and coordinated with other cell-cycle events may lead to insight into the importance of such a substantial periodic transcriptional program.

1.8 Regulators of cell-cycle regulated transcription and the role of CDKs

How is a large and continuous transcriptional program regulated such that it is coordinated with other cell-cycle events? As more and more transcripts were identified as periodic during the cell cycle, focus turned to the regulators that activated or inhibited gene expression – transcription factors (TFs). To identify the regulators that control activation or repression of periodic transcription in

budding yeast, researchers utilized genetic tools and sequence information. Not surprisingly, a number of TFs were found to regulate sets of periodic genes throughout the cell cycle (reviewed in [88-90, 97]). A list of some known TFs involved in cell-cycle regulated transcription are listed in Table 1.1 with relevant information about when each TF is activated and regulation by cyclin/CDKs.

Table 1.1 Transcription factors that are known to play a role in activating or repressing periodic transcription during the cell cycle.

TF	Phase	Function	Representative Target	CDK target?	CDK regulation
SBF	G1/S	activator	*CLN1*	Yes [98]	Inhibitory
MBF	G1/S	activator	*POL1*	Yes [99]	Unknown
Yhp1	G1/S	repressor	*CLN3*	No	N/A
Yox1	G1/S	repressor	*SWI4*	No	N/A
Nrm1	G1/S	co-repressor	N/A	Yes [100]	Unknown
Hcm1	S	activator	*NDD1*	No	N/A
SFF	G2/M	activator	*CLB2*	Yes [101, 102]	Activating
Ace2	M/G1	activator	*NIS1*	Yes [103]	Inhibitory
Swi5	M/G1	activator	*SIC1*	Yes [104]	Inhibitory

The TFs identified possess three striking qualities that suggest potential modes of regulation for the periodic transcription program. First, many of the TFs that play a role in controlling cell-cycle regulated transcription are themselves periodically transcribed (reviewed in [97]). For TFs that act in complexes, at least one TF is periodically expressed. This observation suggests that a portion of genes may be cell-cycle regulated due to the periodic expression of their regulators. Second, cyclin/CDK activity has been found to affect the activity of many of these transcription factors. In these cases, cyclins that are expressed during any of these phases has the capacity to affect TF activity (Table 1.1). Additionally, this regulation can be either be activating or inhibitory depending on the TF (Table 1.1). These findings across multiple TFs imply that cell-cycle regulated gene expression is also modulated by cyclin/CDK activity. Finally, genome-wide binding data have shown that these TFs also bind in the promoter of other TFs shown to regulate periodic gene expression [105-107]. Taken together, these results demonstrated that the TF controlling the last wave of periodic transcription also activates the first TF in the cycle. A TF network, modulated by cyclin/CDK activity, was proposed to account for the periodic nature of the TFs themselves and the entire periodic transcription program ([107], reviewed in [88, 90, 97]). Models for how the TF network is integrated with cyclin/CDK activity and cell-cycle events will be discussed below.

1.8.1 Transcriptional regulators that control periodic transcription

Which TFs are included in this TF network and how is a transcriptional signal transmitted through the cell cycle? Concurrent with passage through Start and the commitment to the cell cycle, the heterodimeric TFs SBF and MBF activate a large program of periodic genes involved in budding, centrosome duplication, and DNA replication. SBF and MBF share a trans-activating subunit, Swi6 [108], and a distinct DNA binding subunit, Swi4 and Mbp1, respectively [109, 110], that activate genes in the above-mentioned processes. Canonical SBF targets are involved in budding and centrosome duplication while MBF targets play a role in DNA replication; however, the extent of functional overlap between SBF and MBF remains unclear [111-113], as *swi4* and *mbp1* single mutants are viable while *swi4 mbp1* double mutants are not [110]. Activation of SBF and MBF centers on a positive feedback loop with G1 cyclin/CDK activity and co-repressor Whi5 [114-117] (Figure 1.3). Cln3/CDK phosphorylates Whi5 that inhibits binding and nuclear localization (Figure 1.3). Cln1 and Cln2 synthesis is activated by SBF and further inhibits Whi5 repression (Figure 1.3). While G1 cyclins are involved in activation, a series of transcriptional repressors and B-type cyclins inactivate SBF and MBF activity. Swi4 periodic gene expression is repressed by SBF targets *YHP1* and *YOX1* [118] and SBF transcriptional activity is repressed by Clb2 [98] (Figure 1.3). The transcriptional activity of MBF is modulated by its target and co-repressor *NRM1* [119] (Figure

1.3). In addition to cyclin/CDK regulation, effectors of the DNA replication checkpoint affect the activity of both SBF and MBF. Rad53 has been shown to phosphorylate Swi6 in order to down regulate *CLN1* and *CLN2* gene expression [120, 121]. Rad53 also has been shown to phosphorylate Nrm1, inhibiting it from binding and repressing MBF activity, resulting in persistent expression of MBF

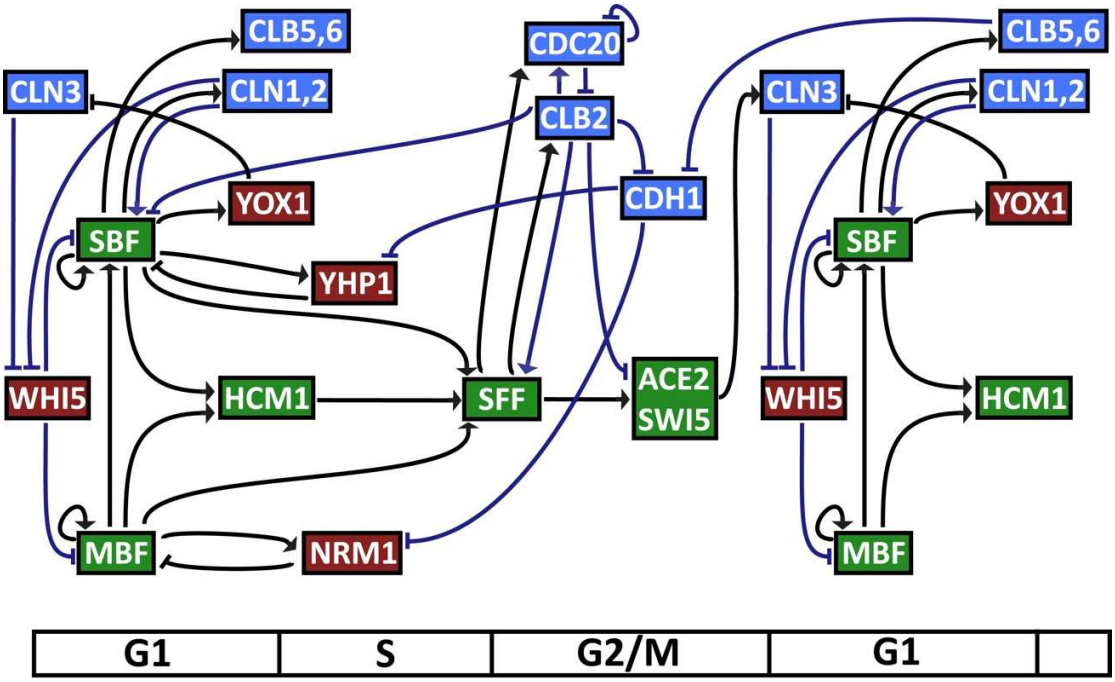

Figure 1.3: A transcription factor network. A number of transcriptional regulators have been shown to affect sets of co-expressed periodic genes. These transcription factors are regulated by one another and by post-transcriptional modifications. All nodes are placed on timeline based on peak time of gene expression in wild-type cells. Green nodes: transcriptional activators; Red nodes: transcriptional repressors; Blue nodes: Post-transcriptional modifiers, cyclins and E3 ubiquitin ligases. Black edges represent the upstream transcription factor binding to the promoter of the downstream node. Blue edges represent post-transcriptional modifications that alter the activity of the transcription factor

targets during the DNA replication checkpoint [122]. Regulation of SBF and MBF is very extensive, as this represents the commitment to the cell cycle.

SBF and MBF transmit a transcriptional signal to activator Hcm1 [113] that is responsible for expression of genes required for chromosome segregation, centrosome dynamics, and budding during late S-phase [80] (Figure 1.3). *HCM1* is periodically transcribed, and its protein levels are also periodic, closely mirroring the behavior of *HCM1* mRNA [80]. Moreover, Hcm1 activates the synthesis of the SBF co-repressor *WHI5* and Swi-five factor (SFF) subunit *NDD1* [80] (Figure 1.3). SFF, a TF complex composed of Fkh1, Fkh2, Ndd1, and Mcm1 [123-127], activates a set of periodic genes referred to as the "*CLB2* cluster" during G2/M phase ([78, 128], reviewed in [88, 90, 97]). SFF activity is modulated through a positive feedback loop with the B-type cyclin Clb2 [98]. SFF activates *CLB2* gene expression [128], which in turn binds CDK and further stimulates components of SFF to increase its transcriptional activity [101, 102, 129] (Figure 1.3). SFF transmits the periodic transcriptional signal by the activating TFs *ACE2* and *SWI5* [128] (Figure 1.3). Ace2 and Swi5 are redundant with each other and activate periodic transcripts involved in the transition between late M phase and the beginning of early G1 of the subsequent cell cycle, such as cell separation. While Ace2 and Swi5 are activated during G2/M, their activity is inhibited by Clb2/CDK-dependent cytoplasmic sequestration [103, 104] (Figure 1.3). After mitosis and the removal of Clb2, Swi5 and Ace2 have been shown to bind to the

promoter of *CLN3* and to activate *CLN3* expression in the next cell cycle [130]. Much of the periodic transcriptional program may be explained by the periodic activity of different cell-cycle TFs. While these TFs are themselves periodically expressed, it is clear that their activity is modulated by cyclin/CDK post-translationally.

1.9 Regulation of periodic transcription: evidence of a transcription factor network

Based on the studies described above, it appears that periodic transcription is regulated during the cell cycle by cyclin/CDK activity through periodically expressed TFs. This mode of regulation is similar to cyclin/CDKs alone initiating other cell-cycle events like budding, centrosome duplication, DNA replication, and chromosome segregation. Could periodic transcription perhaps be just another cell-cycle event that is modulated by oscillations in cyclin/CDK activity?

A series of studies have been carried out to determine the contribution of cyclin/CDK activity on periodic gene expression during the cell cycle. The first study to address this question measured the effect of S-phase and mitotic cyclin/CDKs on periodic transcription by deleting all six of these cyclins (*clb1,2,3,4,5,6*) in budding yeast. These cells are kept alive by the inducible overexpression of Clb1; in the absence of Clb1, they arrest due to the absence of all S-phase and mitotic cyclin/CDK activity which results in their inability to initiate

DNA replication, centrosome duplication or mitosis. However, the arrested cells continue to carry out G1 events, including budding and G1-specific transcription [131]. What happens to the rest of the periodic transcriptional program in these cells? In a synchronous population of early G1 cells lacking all S-phase and mitotic cyclins (synchronized by centrifugal elutriation), global gene expression dynamics were measured by microarray. Strikingly, compared to the expression dynamics of genes normally periodic in wild-type cells, 70% of genes remain periodic in the absence of both S-phase and mitotic cyclins and in the absence of cell-cycle progression with a period very similar to normally-cycling cells [81]. Regardless of the effect cyclin/CDK activity is known to have on periodic transcription, an independent mechanism has the potential to drive a large scale periodic gene expression program.

What kind of mechanism is able to maintain cell-cycle regulated transcription in the absence of any S-phase or mitotic cyclins? Included in the 70% of genes that remain periodic in these cells are many of the TFs involved in modulating periodic transcription throughout the cell cycle [81]. Using these periodic TFs and binding information, Orlando and colleagues were able to construct a TF network that is capable of oscillating *in silico* in the absence of S-phase and mitotic cyclin/CDK activity [81]. This finding led to the proposal that a TF network is capable of functioning as an independent oscillator to modulate periodic transcription ([81], reviewed in [89]). In normally-cycling cells, however,

the cyclin/CDK oscillator and the TF network oscillator are coupled to each other, as cyclins are targets of the TF network and cyclins, when complexed with CDKs, alter the activity of TFs in the network (Table 1.1). Further, it was proposed that this TF network may function as an underlying oscillator responsible for the timing of cell-cycle events, as the TF network is capable of keeping periodic transcription oscillating at a near wild-type period in the absence of cyclin/CDK influence [81]. This proposed model is not substantially different from the cyclin/CDK-centered model of cell-cycle regulated transcription based on previous results. However, rather than cyclin/CDKs acting as the primary regulator of the periodic transcription program, the TF network itself is capable of promoting oscillations in gene expression during the cell cycle.

1.10 Regulation of TF network: alternative models

The observation that periodic transcription can continue in the absence of oscillations in cyclin/CDK activity or of cell-cycle progression is not novel. Other CDK autonomous oscillations have been observed in other cyclin mutants, including budding, centrosome duplication, DNA replication, and Cdc14 nucleolar release [28, 81, 131-133]. Each of these autonomous oscillations follows its own period that is not necessarily synchronized with a wild-type period. However, during a normal cell cycle, all of these events are coordinated with each other in the correct order and with a period that fits with the period of cellular divisions. In cycling cells, what mechanism is responsible for coordinating all of these

autonomous oscillators? A model has been proposed in which oscillations in cyclin/CDKs act as a master oscillator that entrains all other oscillations to the period of the cyclin/CDK oscillator (Figure 1.4; [132], reviewed in [134]). This model is rooted in principles developed in the field of physics and describes how synchrony of oscillations is possible in a concept known as phase locking [135, 136]. In this model, a single oscillator is dominant to other oscillations, forcing other oscillators to alter the period of oscillations to match the main oscillator. In the case of independent cell-cycle oscillators, this model proposes that the cyclin/CDK oscillator functions as the master oscillator, phase locking the other oscillators – budding, DNA replication, centrosome duplication, TF network, and Cdc14 nucleolar release (Figure 1.4). This results in all events occurring once per cell cycle at an observed period dictated by oscillations in cyclin/CDK activity. This model was proposed because these independent oscillations are only reveled in the absence of cyclin/CDK oscillations [28, 81, 131-134]. This model is also very similar to what has previously been proposed, but adds on the layer of how cell-cycle events, including periodic transcription via a TF network, are coupled with each other with the same period.

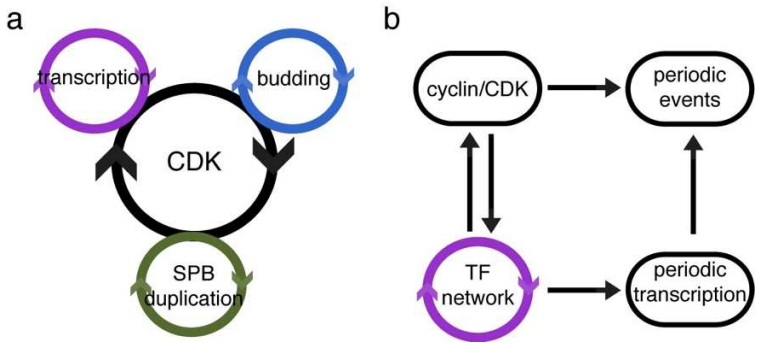

Figure 1.4: Alternative models of cell-cycle regulation. (**a**) One model proposes that oscillations in cyclin/CDK activity control the timing of cell-cycle events, including events with potential autonomous oscillations. (**b**) A second model proposes that a transcription factor network is responsible for the timing of cell-cycle events by regulating the synthesis of periodic genes, cyclins included.

Given this model, questions arise on how the TF network is controlled by oscillations in cyclin/CDK activity. Studies leading to the discovery of an autonomous TF network oscillator showed that oscillations in cell-cycle regulated transcription continue in the absence of all S-phase and mitotic cyclins [81]. Yet, these cells maintain oscillations in G1 cyclin/CDK activity [81, 131]. Could oscillations in G1 cyclin/CDK activity be the driver of cell-cycle regulated transcription in these cells? Previous work has suggested that oscillations in G1 cyclin/CDK activity are responsible for phase locking cells in the commitment period of the cell cycle [137]. In this study, extrinsic periodic pulses of *CLN2* expression were reported to synchronize the cell-cycle period of a population of cycling cells [137]. However, a later study showed that, in the absence of all S-phase and mitotic cyclins and constitutive expression of hyper-stable Cln2, periodic budding continues with a similar period to cells only lacking S-phase and

mitotic cyclins [133]. This finding suggests that the TF network is capable of maintaining oscillations in the absence of oscillations in G1 cyclin/CDK activity and that cells only require the Cln2/CDK activity to initiate budding [133]. Thus, although pulses of extrinsic *CLN2* gene expression is capable of gating cell-cycle progression, it is not necessary for other autonomous oscillations, including budding and periodic transcription.

What contribution do G1 cyclins actually make to regulating periodic transcription? To address this question, global gene expression dynamics were measured by microarray in budding yeast cells lacking a functional CDK using the temperature-sensitive *cdc28-4* allele. Surprisingly, approximately 66% of genes maintain periodicity in the absence of all CDK activity, further suggesting that the bulk of periodic transcription does not require oscillations in cyclin/CDK activity [133]. A TF network could be built with similar structure to TF networks previously constructed and was shown to have capacity to oscillate *in silico* [133]. These observed oscillations occur in the absence of CDK activity. What, then is the role of the TF network in cell-cycle progression? A series of experiments were executed to address this question.

First, it was proposed previously that a TF network may function as an underlying oscillator [81], which controls the proper timing and ordering of cell-cycle events by regulating periodic transcription. If a TF network is the underlying cell-cycle oscillator, changes in expression of the network TFs should change the

period of oscillations. To test this hypothesis, network TFs were perturbed by constitutively overexpressing transcriptional activators or by deleting genes encoding transcriptional repressors in budding yeast upon inhibition of all S-phase and mitotic cyclins. The period of the resulting budding cycles was subsequently measured. The budding cycles are used as a proxy for the function of periodic transcription and the TF network, as budding and periodic transcription both continue in cells lacking all S-phase and mitotic cyclin/CDK activity. In these cells, perturbation of different network TFs altered budding periods, resulting in up to 20% longer or shorter periods compared to cells with unperturbed network TFs [133]. This result suggests that a TF network does function as an oscillator that plays a role in controlling the timing of budding oscillations and periodic transcription.

What then, is the role of cyclin/CDK activity in regulating the timing of cell-cycle events and oscillations? Historical models of cell-cycle regulation postulate that oscillations in cyclin/CDK activity are responsible for the timing of periodic events, including DNA replication, centrosome duplication, and chromosome segregation (reviewed in [26]). As this model was first discovered in early developing embryos and then found to remain conserved in yeast and somatic cells, cell-cycle regulation has centered on fluctuations in cyclin/CDK activity. Oscillations in periodic transcription persist in cells lacking either S-phase and mitotic cyclin/CDK activity or all cyclin/CDK activity, yet critically, cell-cycle

progression is arrested [81, 133]. In yeast and somatic cells, several experiments have shown that cyclin/CDKs have the capacity to alter TF activity. What is the effect that cyclin/CDK feedback acting on TF activity has on transcriptional oscillations? Comparing transcriptional dynamics reveals that the overall amplitude of periodic gene expression dynamics decreases as cyclin/CDK activity is removed [133]. Additionally, the period of transcriptional oscillations also increase with decreasing cyclin/CDK activity [133]. These observations suggest that cyclin/CDK activity plays an important role in regulating the amplitude and period of transcriptional oscillations. In addition to cyclin/CDK feedback on the TF network, cyclins themselves are periodically transcribed (reviewed in [88-90, 97]). Yet in the absence of CDK activity, cell-cycle progression is halted. A study showed that in the absence of all S-phase and mitotic cyclins, except for a single S-phase cyclin, periodic cycles of DNA replication occur together with transcriptional activation of the S-phase cyclin [133]. This observation implies that cyclin/CDK activity also acts as an effector of the TF network oscillator.

This set of experimental results has led to a new model of cell-cycle regulation (Figure 1.4b). A TF network acts as an oscillator that drives the timing of periodic transcription, including cyclins. Cyclins (in complex with CDKs) then feedback onto the TF network via phosphorylation to contribute robust transcriptional oscillations. Phosphorylation of a TF is capable of enhancing or

reducing transactivation of the TF's target genes, thus "fine tuning" the TF network output during the cell cycle. Additionally, cyclin/CDKs also act as effectors of the TF network to trigger cell-cycle events in the proper order (Figure 1.4b). This model is different from previous cell-cycle models in that a TF network, rather than cyclin/CDK activity, acts as the oscillator that keeps the timing and ordering of cell-cycle progression. Although the mechanism by which the timing of cell-cycle oscillations is different, the process by which cell-cycle events are triggered by cyclin/CDKs is not different. The activation of different events in the proper order is dependent both on transcriptional oscillations and cyclin/CDK activity.

While these two models seem very different from each other, they are built around the same results (Figure 1.4). Autonomous transcriptional oscillations are only revealed in the absence of cyclin/CDK activity, although they still occur in normally-cycling cells. Even when perturbing core network TFs, we observe these autonomous oscillations in the absence of S-phase and mitotic cyclin/CDK activity. Much work remains for investigators to do in order to distinguish between these two models. For example, an improved understanding of how a TF network oscillator and its component TFs are regulated during the cell cycle. Additionally, it remains to be determined how a TF network oscillator is coupled to other cell-cycle events and cyclin/CDK activity.

1.11 Concluding remarks

While these two different models seem very different from each other, they are built around the same results. Autonomous transcriptional oscillations are only revealed in the absence of cyclin/CDK activity. Even when perturbing core network TFs, we observe these autonomous oscillations in the absence of S-phase and mitotic cyclin/CDK activity. Several goals remain to distinguish between these two models: to better understand how a TF network oscillator and its component TFs are regulated during the cell cycle, and to delineate how a TF network oscillator is coupled to other control modules, like cyclin/CDK activity and cell-cycle events.

1.11.1 Dissertation outline

This dissertation is composed of four chapters. In chapter 2, I will investigate the two alternative models of cell-cycle regulation, focusing on how a transcription factor network is coupled to cell-cycle progression. Specifically, I will study the roles of CDKs and checkpoints in controlling transcriptional oscillations. Chapter 3 details the computational approaches needed to address the questions posed in Chapter 2, including preprocessing global mRNA measurements and post-processing analytical methods that strive to define the genes that comprise cell-cycle regulated transcription. In chapter 4, I will summarize my work presented in Chapters 2 and 3, and discuss open questions and future work.

Chapter 2 The roles of CDKs and checkpoints in regulating transcriptional oscillations

During normal cell division cycles, periodic events are synchronized. However, in some physiological or experimental conditions, events become uncoupled and autonomous oscillations are revealed. For example, several cell types undergo endocycles, where periodic activation of S-phase is triggered in mitotically-arrested cells. Similarly, we have demonstrated that the cell-cycle transcriptional program in budding yeast has intrinsic oscillatory capabilities, and periodic activation of the program continues in cells depleted for CDK activity. It has been proposed that CDKs comprise the master regulator ensuring that mechanisms with autonomous periodic potential remain coupled during normal cell-cycle progression. Here we show that CDKs do not entrain the autonomous transcriptional program to cell-cycle progression. We also find that activated checkpoint pathways halt the transcriptional oscillator, blocking the periodic activation of cell-cycle regulated genes during a cell-cycle arrest. These findings suggest that checkpoint mechanisms do not act solely to block cell-cycle events, but also to maintain synchrony between the progression of the cell cycle, and the autonomous transcriptional program.

2.1 Introduction

Successful cell divisions depend on proper temporal ordering of periodic cell-cycle events, including DNA replication and segregation. The central oscillator driving periodic events in early embryos is based on the activity of the cyclin/CDK complex [6, 138, 139]. Contemporary models of cell-cycle control in yeast and metazoan cells also center on fluctuations in mitotic CDK activity, but include additional cyclins, feedback loops, and transcription [5]. However, we have challenged these models by demonstrating that cell-cycle events, including budding, centrosome duplication, DNA replication, and periodic transcription continue, even when budding yeast cells were arrested in G1 by depletion of CDK activity [81, 131, 133]. An autonomously oscillating transcription factor network was proposed to function independently of CDKs and cell-cycle progression to control the periodic transcriptional program[81, 89, 133]. However, transcriptional network oscillations are coupled to oscillations in CDK, as cyclins are periodically transcribed and CDKs, in turn, affect the activity of network transcription factors[81, 89, 133]. A recent study has shown that when cell-cycle progression is arrested in M phase with varying levels of mitotic cyclin Clb2, autonomous oscillations in Cdc14 nucleolar release are observed[132], leading to the proposal that CDK functions as a master cell-cycle oscillator that couples other observed autonomous oscillations in order to ensure synchronous activation of cell-cycle events [132, 134]. This model is consistent with previous

findings, as autonomous transcriptional network oscillations may only be revealed in the absence of CDK oscillations (CDK “off”). What role, if any, does CDK play in regulating the timing of autonomous transcriptional network oscillations?

Mitotic cyclin, Clb2, when bound to CDK is known to both activate and inhibit several waves of transcription (Figure 2.1). First, Clb2/CDK is involved in a positive feedback loop with SFF to further activate its own synthesis (Figure 2.1) [101, 102, 129].

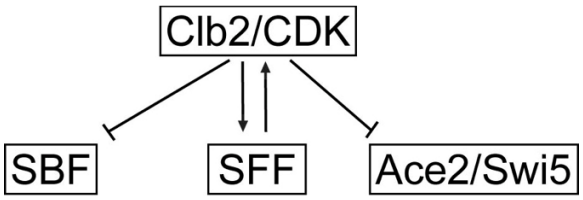

Figure 2.1: Clb2/CDK activity alters the activity of transcription factors during the cell cycle.

Further, Clb2/CDK activity inhibits G1/S transcription by repressing SBF activity[98] and inhibits M/G1 transcription by sequestering redundant transcription factors Swi5 and Ace2 in the nucleus (Figure 2.1) [103, 104]. In the absence of S-phase and mitotic cyclin/CDK activity, the positive feedback and repression is lost in addition to a cell-cycle arrest in G1, but periodic transcription continues[81, 133]. What is the effect of persistent mitotic cyclin/CDK activity on periodic transcription and the autonomous transcription factor network? Degradation of Clb2 during mitosis triggers chromosome segregation at the

metaphase-to-anaphase transition[140, 141]. How does a cell-cycle arrest in G2/M with persistent mitotic cyclin/CDK activity (CDK "on") in *Saccharomyces cerevisiae* affect periodic transcription and the autonomous transcription factor network? Compared to cells lacking CDK activity (*clb1,2,3,4,5,6*; CDK "off"), substantive change in the dynamics of periodic transcription is expected in cells arrested with the CDK oscillation is locked in the "on" state.

2.2 Transcriptional oscillations are maintained with persistent mitotic cyclin/CDK activity

2.2.1 CDK "on" cells arrest with a single bud, a short spindle, and persistent Clb2 protein levels

To arrest yeast cells with the CDK oscillation locked in the "on" state, we used a strain in which the anaphase promoting complex (APC) activator, is conditionally expressed from a modified *GAL1* promoter (P_{GALL}-*CDC20*) in a *cdc20*Δ background[130]. When cells are shifted from galactose medium to glucose medium, Cdc20 is rapidly depleted, arresting cells at the metapahase-to-anaphase transition with persistent levels of Clb2 protein (Figure 2.2) and Clb2/CDK activity[140, 141]. A synchronous population of P_{GALL}-*CDC20 cdc20*Δ cells in early G1 was collected by centrifugal elutriation, suspended in dextrose growth medium, and then aliquots of cells were collected at 20-minute intervals for 300 or 360 minutes (two experimental replicates). Genome-wide transcript levels were then assayed at each time point by microarray. Cell-cycle

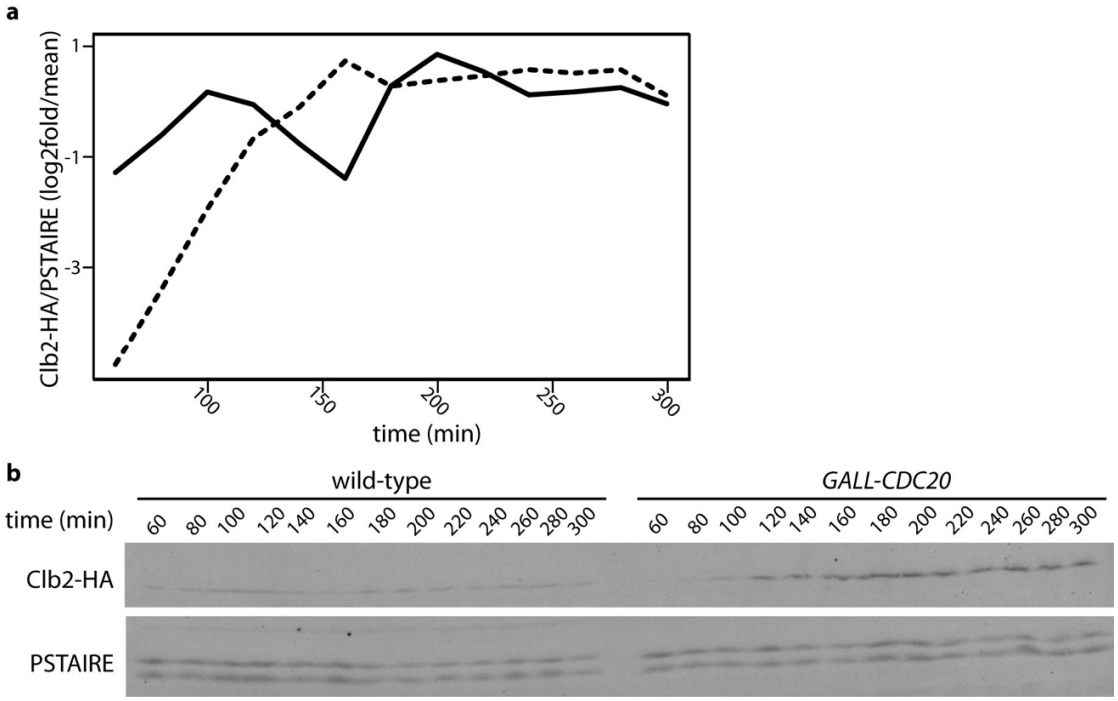

Figure 2.2: Clb2 protein dynamics in *GALL-CDC20* and wild-type cells. Dynamics of Clb2 protein levels were measured in *GALL-CDC20* and wild-type cells by tagging Clb2 with an HA tag. Quantified levels of Clb2 protein normalized to PSTAIRE and $\log_2$-transformed relative to the mean (a). Solid line, wild-type cells; dashed line, *GALL-CDC20* cells. Representative blot of three replicate experiments of Clb2-HA levels in wild-type or *GALL-CDC20* cells (b).

progression and subsequent arrest was monitored by budding and spindle length to ensure a cell-cycle arrest (Figure 2.3). Results from two independent replicates were highly reproducible, with an r^2 value of 0.98 (Figure 2.4).

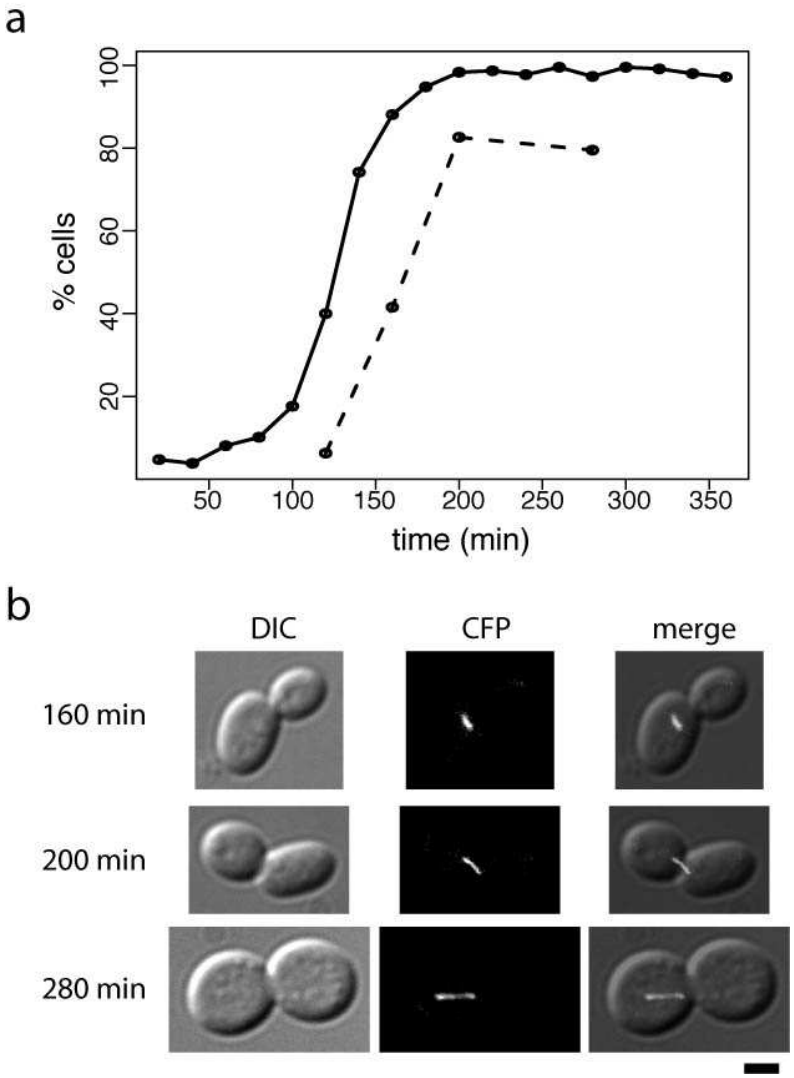

Figure 2.3: Cell-cycle arrest phenotype of P_{GALL}-*CDC20* cells. Dynamics of budding (solid line) and short spindle formation (dashed line) in G1-synchronized P_{GALL}-*CDC20* cells released into dextrose-containing medium (a). These cells terminally arrest with a large bud (left, right panel) and short spindle (middle, right panel) Bar is 5µm (b).

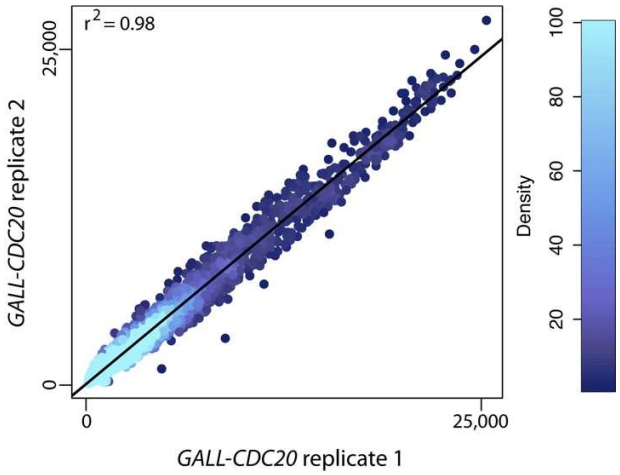

Figure 2.4: Reproducibility of CDK "on" repliates. For each probe, the absolute mean expression value/1000 was calculated and plotted (arbitrary units) for each replicate. Coloring of each dot indicates the density of points surrounding the probe in a square with length 500 centered on that point. These data were fitted to a linear model (black line) and the corresponding r^2 value is given.

2.2.2 TFs targeted by CDK activity affect their targets' gene expression in a predictable manner

As expected, the dynamic transcript behaviors of SBF-, SFF-, Swi5-, and Ace2-regulated genes in CDK "on" cells, CDK "off" cells, and wild-type cells are consistent with the known regulatory interactions between network transcription factors and Clb2/Cdk1 (Figure 2.5). Transcripts from SBF-regulated G1/S genes are repressed after the first cycle of expression in CDK "on" cells and are expressed at elevated levels, but continue to oscillate, in CDK "off" cells (Figure 2.5a). Many transcripts from G2/M genes regulated by SFF are expressed at persistent levels in CDK "on" cells, likely due to positive feedback with Clb2[102] (Figure 2.5b). The positive feedback between Clb2/CDK and SFF results in

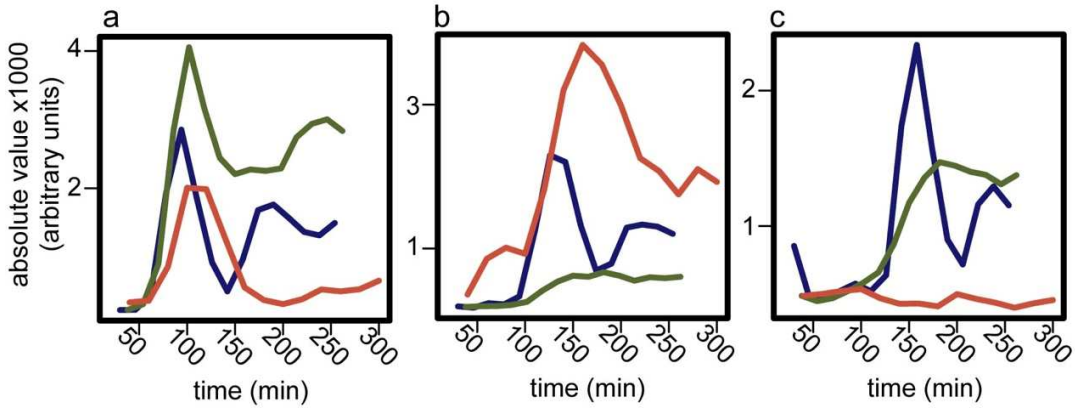

Figure 2.5: Persistent Clb2/CDK activity affects known CDK-regulated genes. Absolute mRNA levels (dChip-normalized Affymetrix intensity units/1000) are shown for periodic genes in CDK-on (red), CDK-off (green), and wild-type (blue) cells. *CLN2*, SBF target (**a**), *SWI5*, SFF target (**b**), and *SIC1*, Ace2/Swi5 target (**c**).

persistent expression, but not at elevated levels compared to normally-cycling cells. This suggests the possibility of some other unknown mechanism that modulates the positive feedback loop. M/G1 genes regulated by the transcription factors Ace2 and Swi5 are not expressed in CDK “on” cells because Clb2-dependent phosphorylation of Swi5 and Ace2 sequesters the transcription factors in the cytoplasm[103, 104] (Figure 2.5c). While we find that Clb2/CDK activity does regulate the gene expression dynamics of expected periodic genes, what effect does this have on the global periodic transcription program and the autonomous transcription factor network?

2.2.3 Methods in identifying periodic genes in wild-type cells

We ran the wild-type gene expression dynamics through the deLichtenberg algorithm using the average cell-cycle period. At a p-value cutoff of $p \leq 0.2$, we identified 1912 wild-type periodic genes. Of the 1275 periodic genes from Orlando et al[81], 1010 genes were also classified as oscillating at the average cell-cycle period. In addition to a large overlap with Orlando et al[81], the averaged cell-cycle period adds 902 periodic genes.

Additionally, periodicity of wild-type gene expression dynamics was scored using the Lomb-Scargle algorithm across a range of periods from 50 minutes to 300 minutes [142-144]. We chose a period of 97 minutes with a ten-minute window around that average cell-cycle period, and a p-value cutoff of $p \leq 0.5$. This results in 991 periodic genes in wild-type cells.

Using genes only identified as periodic in both the deLichtenberg and Lomb-Scargle analyses, we generated a high-confidence wild-type periodic gene list. This results in a wild-type periodic gene list containing 856 genes (Figure 2.6).

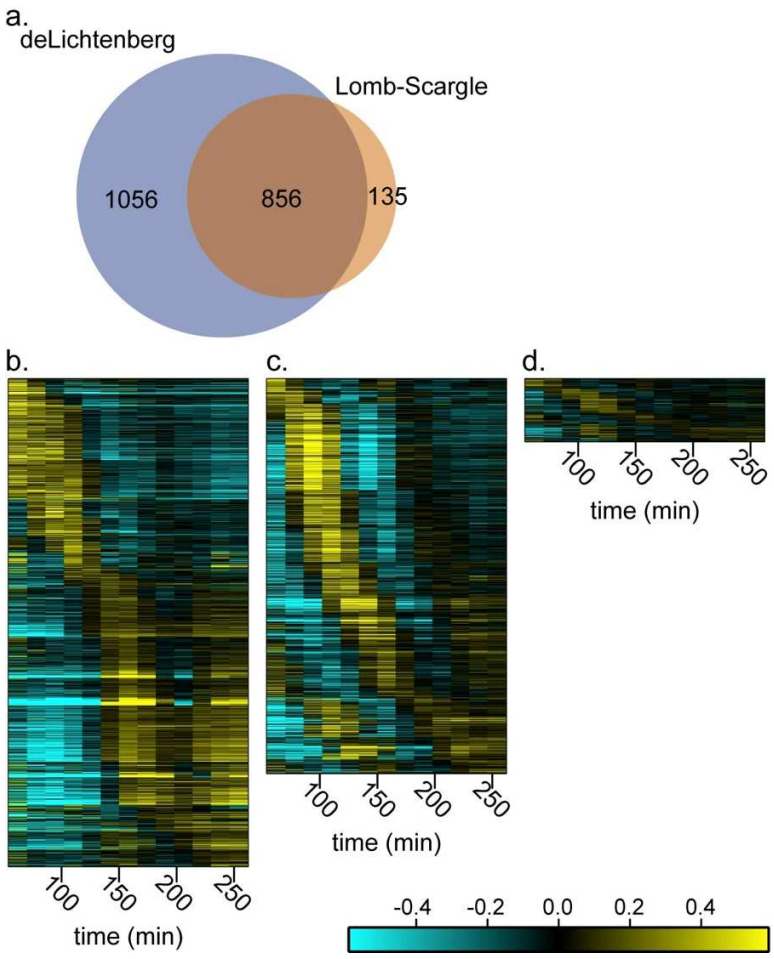

Figure 2.6: Wild-type restrictive periodic gene list. Venn diagram of the resulting periodic gene lists from the deLichtenberg algorithm at the CLOCCS average cell-cycle period and the Lomb-Scargle algorithm at a 10 minute period range centered around the CLOCCS average cell-cycle period and a $p \leq 0.5$ cutoff (a). Heat maps showing the mRNA levels of genes that are periodic only in deLichtenberg (b), that are periodic in both lists (c), and that are periodic only in Lomb-Scargle (d). Each heat map displays a different set of genes. Transcript levels are depicted as $\log_2$-fold change relative to the mean expression

2.2.4 Methods in identifying periodic genes in CDK “on” cells

In previous studies, we have used Pearson correlation as a basis to identify wild-type cell-cycle regulated genes that maintain periodicity in cells depleted for CDK activity (*clb1,2,3,4,5,6* and *cdc28-4*)[81, 133]. While Pearson correlation has successfully identified genes that look the same across two experimental conditions, this method does not directly address what genes in these conditions are periodic. Additionally, Pearson correlation methods score gene expression dynamics as similar even if only one cycle of periodic expression agrees and the remaining dynamics do not match. Utilizing this approach may lead to incorrect classification of genes that remain periodic in CDK “on” (Cdc20-depleted) cells.

To identify oscillating gene expression dynamics in non-wild-type cells, each non-wild-type replicate was run through Lomb-Scargle at the same period range as wild-type [142-144]. We did not use deLichtenberg to analyze these data sets due to the permutation-based method to score periodicity and amplitude. This method makes comparing outputs difficult to do since permuting the data will be different with every experiment.

Cells arrested with persistent mitotic CDK activity do not display any observable periodic oscillations (P_{GALL}-CDC20; Figure 2.3). To generating a periodic gene list for these datasets, a ten-minute period range centered on 148 minutes was selected. Lomb-Scargle classifies the largest number of genes at

this period and p-value cutoff. To be consistent with the period range chosen for the wild-type gene lists, a ten-minute period range with a p-value cutoff of $p \leq 0.5$ was selected to generate a periodic gene list.

2.2.5 A TF network maintains oscillations in CDK "on" cells

Surprisingly, a large number of transcripts continued to oscillate in CDK "on" (Cdc20-depeleted) cells (Figure 2.7a). Starting with a high confidence set of 856 genes shown to be periodic in wild-type cells (Figure 2.6), we found that 206 cell-cycle-regulated transcripts continued to oscillate with dynamics similar to wild-type cells (Figure 2.7c). Furthermore, of the 206 transcripts that oscillate in CDK "on" cells, approximately 40% also oscillate in CDK "off" cells (Figure 2.7a and b). Strikingly, MCMC analyses[133] indicate the period of oscillations is nearly identical between the CDK "on" cells and the CDK "off" cells (Figure 2.7d,e, and f, red and green lines). Taken together, these results indicate that the global oscillatory functions of the transcription factor network are largely unaffected by the state of the CDK activity.

These observations suggest that while arresting cell-cycle progression with stable mitotic CDK activity affects periodic transcription locally, the transcription factor network oscillator still functions to maintain a global pattern of the periodic transcription program. How does the transcription factor network continue in cells with persistent B-cyclin/CDK activity regardless of CDK influence at several cell-cycle regulated transcription factors? Although SBF-

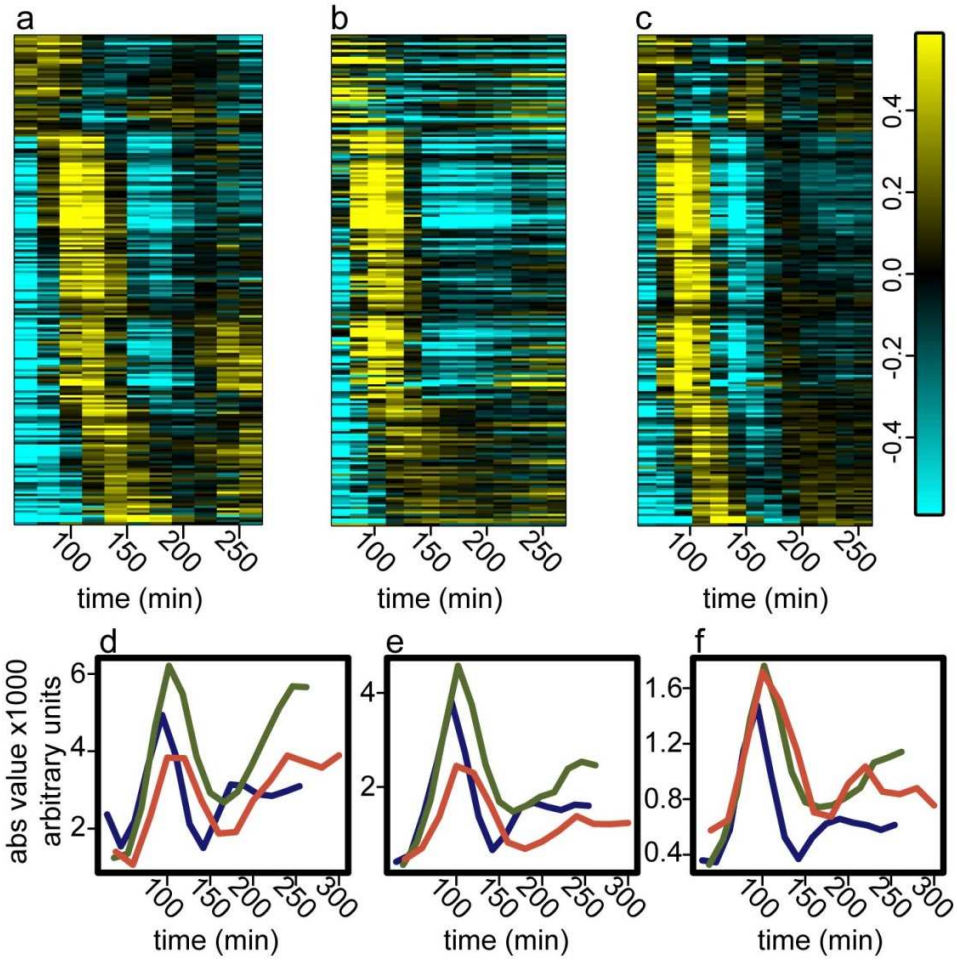

Figure 2.7: Dynamics of periodic transcripts in arrested cells with CDK-on, CDK-off, and cycling wild-type cells. Heat map shows mRNA levels of genes in cells arrested with CDK-on (cdc 20 mutant) (a). Heat maps show behaviors of the same genes in arrested cells with CDK-off (cyclin mutant cells (b), and in normally cycling wild-type (c) cells. The same order of genes is found in all conditions (Supplementary Table 2). Transcript levels are depicted as $\log_2$-fold change relative to the mean expression. Individual gene expression profiles of *RFA2* (d), *PDS5* (e), and *SPH1* (f). Blue, wild-type; green, *clb1,2,3,4,5,6*; red, P_{GALL}-*CDC20*.

mediated transcription is inhibited in the CDK "on" cells, oscillation appears to continue through a parallel branch of then network containing the related transcription factor complex, MBF. While SBF activity is inhibited by Clb2, MBF is

not. Rather, MBF-mediated transcription is normally inactivated through a negative feedback loop with its own target, Nrm1[119]. MBF and SBF have distinct but overlapping sets of target genes, and genetic studies indicate they are functionally redundant[111-113]. In cells Cdc20-depleted cells, we find that targets of all transcription factors of the network oscillator[81, 133] except for SBF, Ace2, and Swi5 maintain oscillatory behavior. Thus, the oscillatory function of the network is likely supported by redundant sets of transcription factors. Consistent with this finding, it has been shown that the perturbation of several transcription factors within the network oscillator does not abrogate transcriptional oscillations[133].

2.3 Role of checkpoints in coupling a TF network to cell-cycle events

Thus far, we have shown that a transcription factor network oscillator can be uncoupled from cell-cycle progression and CDK oscillations (Fig. 2.7) [81, 133]. How then is synchrony maintained between transcription factor network oscillations and cell-cycle progression? In normally cycling cells, checkpoint pathways monitor cell-cycle progression to ensure that events are completed with the proper temporal order [30]. In response to perturbation of cell-cycle events, a checkpoint pathway is activated and subsequently halts cell-cycle progression and blocks CDK oscillations until events are completed with fidelity.

Thus, we hypothesize that checkpoints are likely candidates for synchronizing the transcription factor network oscillator with cell-cycle progression.

2.3.1 Two cell-cycle checkpoints arrest cell-cycle progression by inhibiting chromosome segregation

To determine whether cell-cycle checkpoints can control the dynamics of the transcription factor network oscillator, we measured global transcript dynamics in synchronized populations of cells arresting at either the DNA replication checkpoint or spindle assembly checkpoint. The DNA replication checkpoint was triggered using a temperature sensitive allele of thymidylate kinase gene,*cdc8*, and the spindle checkpoint was triggered by constitutive over-expression of a mutant allele of the kinetochore protein, *CSE4* (P_{GAL1}-*cse4-353* [145]). Checkpoint-mediated cell-cycle arrest was monitored by measuring budding index, and either DNA content or spindle length (Figure 2.8). Genome-wide transcript levels were measured by microarray. Results from two independent replicates were highly reproducible for the DNA replication and spindle assembly checkpoints, with an r^2 value of 0.994 and 0.933, respectively (Figure 2.9).

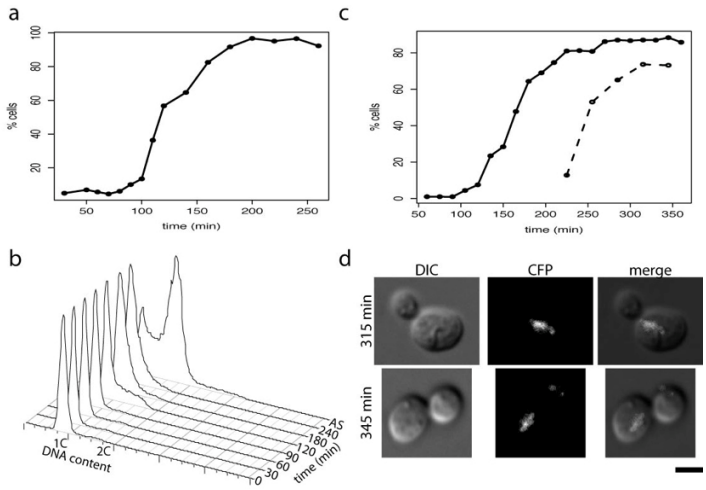

Figure 2.8: Cell-cycle arrest phenotype of checkpoint-arrested cells. Dynamics of budding in G1-synchronized *cdc8*^{ts} cells released into the restrictive temperature (30°) (a). DNA content of *cdc8*^{ts} cells (b). Dynamics of budding (solid) and short spindle formation (dashed) of *GAL-cse4-353* cells released into galactose medium to trigger the spindle assembly checkpoint (c). These cells terminally arrest with a large bud (left, right panel) and short spindle (middle, right panel) (d).

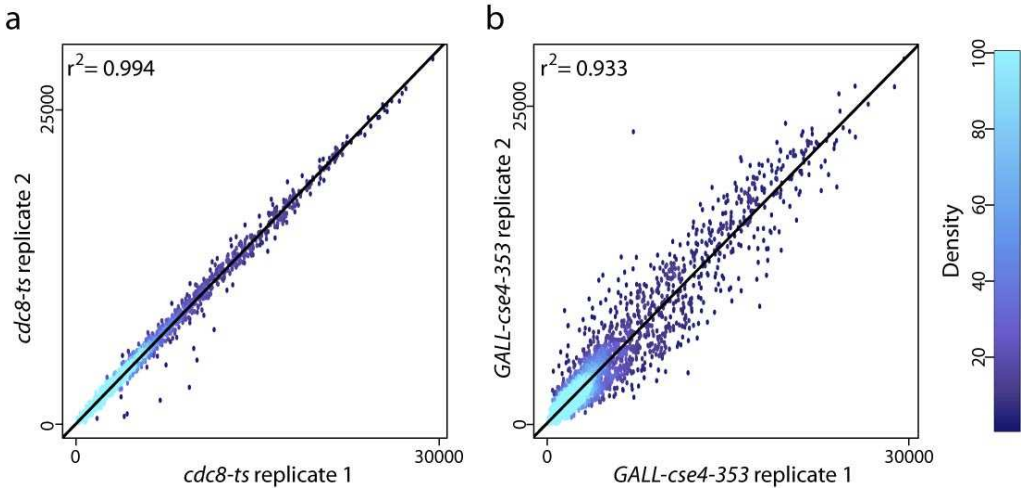

Figure 2.9: Reproducibility of *cdc8*^{ts} and *GAL-cse4-353* replicates. The reproducibility of *cdc8*^{ts} (a) and *GAL-cse4-353* (b) replicate experiments. For each probe, the absolute mean expression value/1000 was calculated and plotted (arbitrary units) for each replicate. Coloring of each dot indicates the density of points surrounding the probe in a square with length 500 centered on that point. These data were fitted to a linear model (black line) and the corresponding r^2 value is given.

2.3.2 Cell-cycle checkpoints halt the periodic transcription program

Transcript dynamics for the vast majority of the genes in high-confidence periodic gene set (856 genes) came to a halt after the activation of both checkpoints (Figure 2.10b and c).

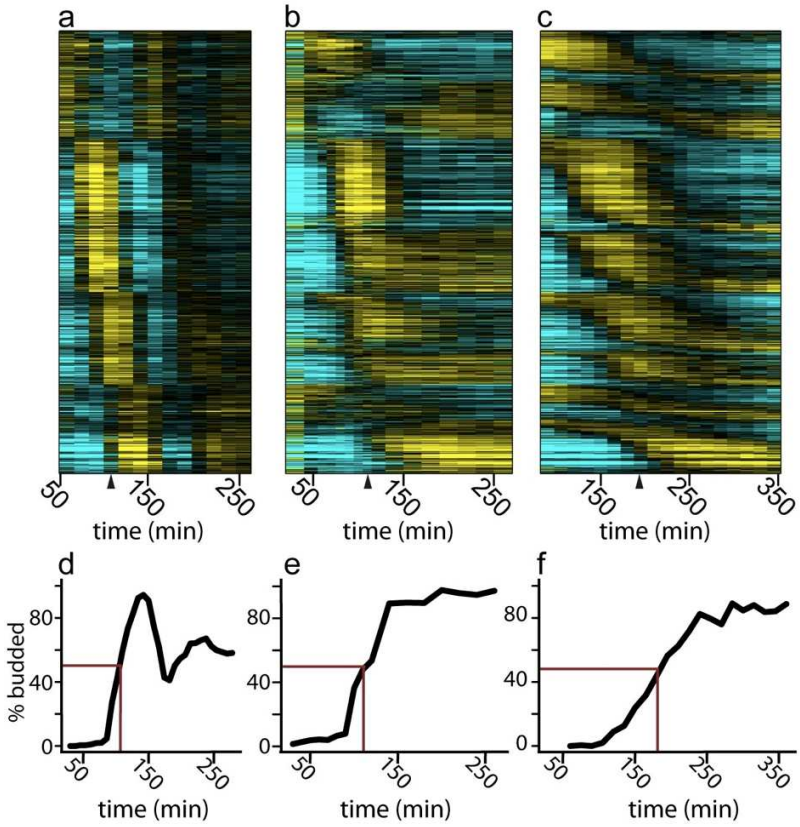

Figure 2.10: Dynamics of periodic transcripts during the DNA replication and spindle assembly checkpoints. Heat maps showing mRNA levels of all wild-type periodic genes in cycling wild-type cells (a), in DNA replication checkpoint arrested (*cdc8*^{ts}) cells (b) and spindle assembly checkpoint arrested (*GAL1-cse4-353*) cells(c). The same order of genes is found in all conditions (Supplementary Table 1). Transcript levels are depicted as $\log_2$-fold change relative to the mean expression. Black arrow indicates time at which 50% of population is budded. Budding indices of wild-type (d), *cdc8*^{ts} (e), and *GAL1-cse4-353* cells (f). Black line, % budded; red line, 50% one bud.

Although clustering analyses did reveal approximately 15 genes that appear to continue to oscillate in both experimental replicates (Figure 2.11 and Figure 12). By visual inspection, we identified two clusters in the DNA replication checkpoint (clusters 7 and 18, Figure 2.11) and 2 clusters in the spindle assembly checkpoint (clusters 6 and 14, Figure 2.12) that may exhibit periodic behaviors. Upon visual inspection of the individual genes included in these clusters, only approximately 34 genes remain periodic during the DNA replication checkpoint and 18 genes remain periodic during the spindle assembly checkpoint.

Previous genomic studies utilizing non-synchronized cells identified only a handful of transcripts that appeared to be regulated by the DNA replication or damage checkpoints [59], yet we observe that nearly the entire cell-cycle-regulated transcriptional program appears to be controlled by the DNA replication checkpoint. Our ability to identify these broad changes in transcriptional behaviors likely reflects the use of synchronous populations of cells and high-density sampling of the time-series. A recent study using synchronous cells and lower density sampling also indicated that the expression of large clusters of cell-cycle regulated genes would be controlled by the DNA replication checkpoint[122]. Furthermore, our observations reveal that the spindle checkpoint also controls cell-cycle regulated gene expression, and utilizes pathways distinct from the DNA replication checkpoint (Figure 2.10).

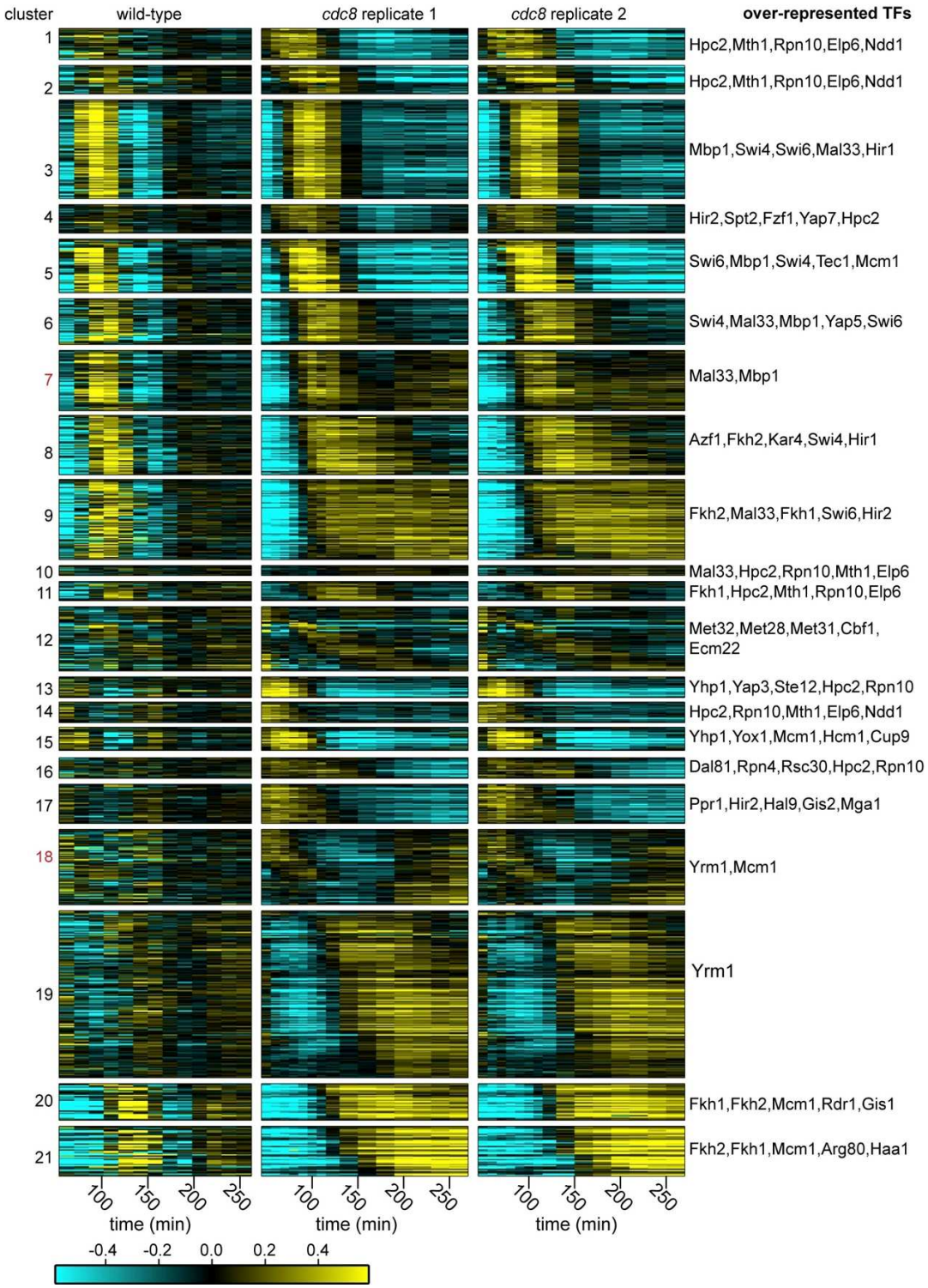
cluster
wild-type
cdc8 replicate 1
cdc8 replicate 2
over-represented TFs
1
Hpc2,Mth1,Rpn10,Elp6,Ndd1
2
Hpc2,Mth1,Rpn10,Elp6,Ndd1
3
Mbp1,Swi4,Swi6,Mal33,Hir1
4
Hir2,Spt2,Fzf1,Yap7,Hpc2
5
Swi6,Mbp1,Swi4,Tec1,Mcm1
6
Swi4,Mal33,Mbp1,Yap5,Swi6
7
Mal33,Mbp1
8
Azf1,Fkh2,Kar4,Swi4,Hir1
9
Fkh2,Mal33,Fkh1,Swi6,Hir2
10
Mal33,Hpc2,Rpn10,Mth1,Elp6
11
Fkh1,Hpc2,Mth1,Rpn10,Elp6
12
Met32,Met28,Met31,Cbf1,
Ecm22
13
Yhp1,Yap3,Ste12,Hpc2,Rpn10
14
Hpc2,Rpn10,Mth1,Elp6,Ndd1
15
Yhp1,Yox1,Mcm1,Hcm1,Cup9
16
Dal81,Rpn4,Rsc30,Hpc2,Rpn10
17
Ppr1,Hir2,Hal9,Gis2,Mga1
18
Yrm1,Mcm1
19
Yrm1
20
Fkh1,Fkh2,Mcm1,Rdr1,Gis1
21
Fkh2,Fkh1,Mcm1,Arg80,Haa1
100
150
200
250
time (min)
-0.4
-0.2
0.0
0.2
0.4

Figure 2.11 Distinguishing different gene expression dynamics during the DNA replication checkpoint. Wild-type periodic genes were clustered by affinity propagation using the first DNA replication checkpoint replicate (*cdc8*^{ts}) expression dynamics. Heat maps showing the mRNA levels of clusters in wild-type (left) and in DNA replication checkpoint-arrested cells (middle,right). Ordering is the same across conditions and replicates. Transcript levels are depicted as $\log_2$-fold change relative to the mean expression. Up to five over-represented transcription factors for each cluster are shown.

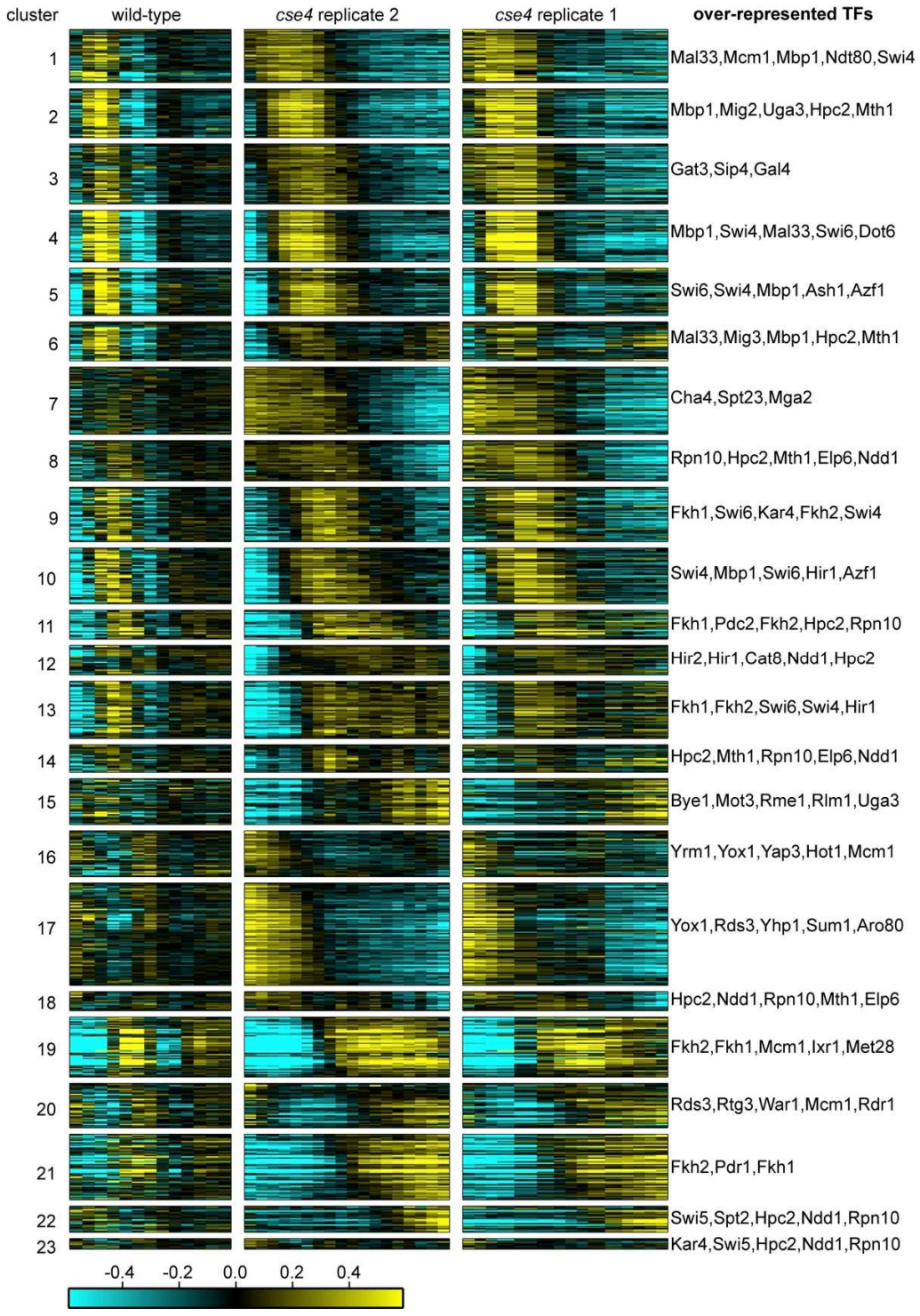
cluster
wild-type
cse4 replicate 2
cse4 replicate 1
over-represented TFs
1
Mal33,Mcm1,Mbp1,Ndt80,Swi4
2
Mbp1,Mig2,Uga3,Hpc2,Mth1
3
Gat3,Sip4,Gal4
4
Mbp1,Swi4,Mal33,Swi6,Dot6
5
Swi6,Swi4,Mbp1,Ash1,Azf1
6
Mal33,Mig3,Mbp1,Hpc2,Mth1
7
Cha4,Spt23,Mga2
8
Rpn10,Hpc2,Mth1,Elp6,Ndd1
9
Fkh1,Swi6,Kar4,Fkh2,Swi4
10
Swi4,Mbp1,Swi6,Hir1,Azf1
11
Fkh1,Pdc2,Fkh2,Hpc2,Rpn10
12
Hir2,Hir1,Cat8,Ndd1,Hpc2
13
Fkh1,Fkh2,Swi6,Swi4,Hir1
14
Hpc2,Mth1,Rpn10,Elp6,Ndd1
15
Bye1,Mot3,Rme1,Rlm1,Uga3
16
Yrm1,Yox1,Yap3,Hot1,Mcm1
17
Yox1,Rds3,Yhp1,Sum1,Aro80
18
Hpc2,Ndd1,Rpn10,Mth1,Elp6
19
Fkh2,Fkh1,Mcm1,Ixr1,Met28
20
Rds3,Rtg3,War1,Mcm1,Rdr1
21
Fkh2,Pdr1,Fkh1
22
Swi5,Spt2,Hpc2,Ndd1,Rpn10
23
Kar4,Swi5,Hpc2,Ndd1,Rpn10
-0.4
-0.2
0.0
0.2
0.4

Figure 2.12: Distinguishing different gene expression dynamics during the spindle assembly checkpoint. Wild-type periodic genes were clustered by affinity propagation using the second spindle assembly checkpoint replicate ($P_{GAL1\text{-}10}$-*cse4-353*) expression dynamics. Heat maps showing the mRNA levels of clusters in wild-type (left) and in spindle assembly checkpoint-arrested cells (middle,right). Ordering is the same across conditions and replicates. Transcript levels are depicted as $\log_2$-fold change relative to the mean expression. Up to five over-represented transcription factors for each cluster are shown.

2.3.3 Cell-cycle checkpoints arrest periodic transcription by affecting network TFs

Clues to the mechanisms by which checkpoints control the transcriptional network are evident when examining specific regulons of co-regulated genes. For example, periodic genes regulated by the transcription factor complex MBF are persistently expressed during the DNA replication checkpoint while these same genes are only expressed for one cycle and subsequently repressed during the spindle assembly checkpoint (Figure 2.13b, and 2.11-12). Persistent expression of MBF targets during the DNA replication checkpoint is consistent with the behavior of MBF targets in *Schizosaccharomyces pombe*[146]. Further studies in *S. pombe* indicate checkpoint control of MBF activity is mediated by Rad53-dependent regulation of the MBF activator Cdc10[146] and the co-repressor Nrm1[147] . Recent reports indicate Rad53 inactivates Nrm1 as part of the DNA replication checkpoint in *S. cerevisiae* [122]. Down-regulation of SBF targets in both DNA replication checkpoint- and spindle assembly checkpoint-arrested cells suggests that stabilization of Clb2 by inhibition of APC is responsible for control

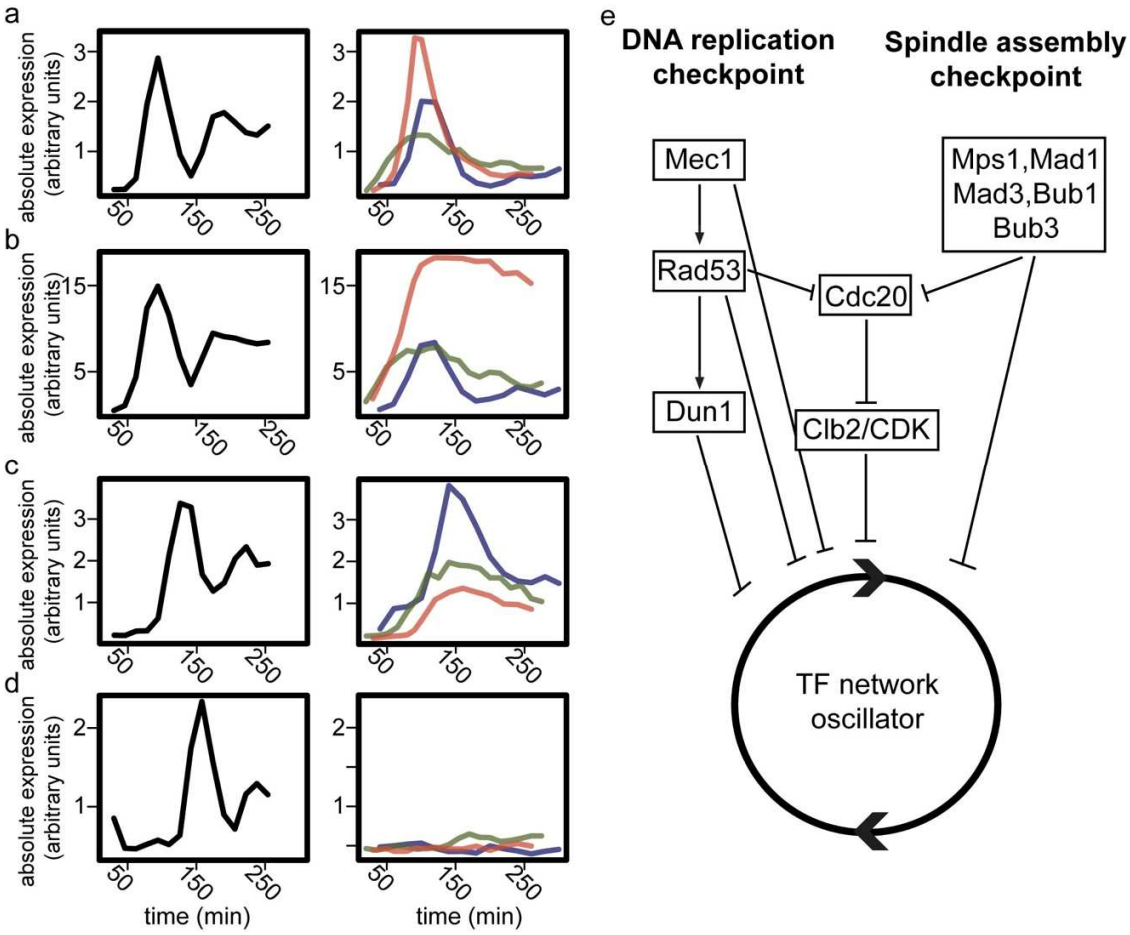

Figure 2.13: Complex regulation of transcriptional oscillator during cell-cycle checkpoints. Absolute mRNA levels (dChip-normalized Affymetrix intensity units/1000) are shown for periodic genes in normally-cycling (blue), DNA replication checkpoint-arrested (red), and spindle assembly checkpoint-arrested (black) cells (a-d). mRNA curves from *cdc8*^{ts} and $P_{GAL1\text{-}10}$-*cse4-353* cells were aligned by 50% budding. *CLN2*, SBF target (a), *RNR1*, MBF target (b), *CLB2*, SFF target (c), and *SIC1*, Ace2/Swi5 target (d). Model proposing potential interactions between checkpoint-specific effectors and the transcriptional oscillator that may also function to arrest cell-cycle regulated transcription (e).

of this regulon (Figure 2.13a, and Figure 2.11-12). Similar transcript dynamics are observed for the two checkpoints in both the SFF- and Swi5/Ace2-regulated clusters (Figure 2.13c,d, and Figure 2.11-12). These observations suggest that

effector pathways distinct for each checkpoint are responsible for regulating the transcription factors within the network oscillator. Rad53, Cdc20, and Dun1 are effector proteins known to control transcription in response to activation of the DNA replication checkpoint (Fig. 2.13e). However, for the spindle assembly checkpoint, it is unclear which effector pathways, in addition to Cdc20, may be controlling the transcription network oscillator (Fig. 2.13e). Identifying the checkpoint effector pathways and their cognate targets in the transcription network oscillator will be the next step towards understanding how checkpoint pathways arrest the periodic activation of cell-cycle transcription. Taken together, our findings indicate that checkpoint pathways regulate the transcription network in order to maintain synchrony between cell-cycle progression and the transcription network oscillator.

2.4 Discussion

The historic model for cell-cycle regulation is centered on two regulatory modules: (1) the CDK oscillator driving events and some periodic transcription and (2) checkpoints arresting events and the CDK oscillator. However, recent studies have proposed a third regulatory module, a transcription factor network, that functions as a the underlying oscillator driving the timing of cell-cycle events with CDK activity serving as an effector of the network[133]. Here, we connect the checkpoint regulatory module to the transcription factor network oscillator

module and provide a mechanism by which the autonomous transcription factor network oscillator can be coordinated with cell-cycle progression.

2.5 Experimental details

2.5.1 Strains and synchronization

Wild-type and all mutant strains of *S. cerevisiae* are derivatives of BF264-15Dau and were constructed by standard yeast methods. A description of all yeast strains and plasmids used in this study are outlined in Table 1. Standard growth conditions were used. Cell synchronization methods were previously described[81, 133].

Table 2.1 Strain list

Strain	Relevant Genotype	Reference
15Da^{-}	*MATa; bar1;ade1;his2;leu2-3112;trp1-1;ura3Δ3*	this lab
SBY1621	*MATa; bar1;cdc20::LEU2;Pgall-CDC20::ADE2;CFP-TUB1::URA3*	[130]
SBY353	*MATa; BAR1;cdc8^{ts}*	this lab
SBY1592	*MATa;Pgal-cse4-353::TRP1;CFP-TUB1::URA3*	[145]
SBY1621	*MATa;bar1;cdc20::LEU2;Pgall-CDC20::ADE2;CFP-TUB1::URA3;CLB2-HA:kanMX6*	this lab
SBY1258	*MATa;bar1;CLB2-HA:kanMX6*	this lab

2.5.2 RNA isolation and microarray analysis

For all global gene expression studies, total RNA was isolated at time intervals by methods described previously[131]. RNA was purified and concentrated using the RNAeasy MinElute Cleanup Kit (QIAGEN). mRNA amplification and fluorescent labeling was done using either the GeneChip One Cycle Labeling (Affymetrix) or the Ambion MessageAmp Premier kit (Ambion Biosystems). Labeled cDNA was hybridized to Yeast 2.0 Expression arrays (Affymetrix) and image collection was carried out by the Duke Microarray Core Facility (http://microarray.genome.duke.edu/) using standard Affymetrix protocols.

CEL files from the Affymetrix Yeast 2.0 oligonucleotide arrays generated from this study (*cdc20Δ;*P_{GALL}*-CDC20*, *cdc8*^{ts}, and P_{GAL}*-cse4-353*) and those from previous studies (wild-type and *clb1,2,3,4,5,6*)[81] were normalized and summarized using a modified version of the dChip[148] method from the `affy` package (v. 1.32.1) in Bioconductor[149] within the R statistical programming environment[143]. The probes specific to *S. pombe* were removed using the s_cerevisiae.msk file from Affymetrix before normalizing the CEL files. The new command used to normalize and summarize the CEL files was `expresso2(inputdata, normalize.method= ”invariantset”, bgcorrect.method= ”none”, pmcorrect.method= ”pmonly”, summary.method =”liwong”, verbose= TRUE)`. Consistent with the

original dChip algorithm, this command outputs a measure of the absolute transcript level for each probe in arbitrary units.

2.5.3 Protein isolation and immunoblotting

Cell lysates were subjected to SDS-PAGE and immunoblotting with the following antibodies. For Clb2 protein measurements (Supplementary Figure 1): mouse anti-HA (Roche Diagnostics), mouse anti-PSTAIRE (Abcom, Inc.), and IRDye 800 conjugated goat anti-mouse (Li-Cor Biosciences). These membranes were analyzed with a Li-Cor Odyssey Infrared Imaging System (Li-Cor Biosciences). Signal was quantified using ImageJ1.41o (National Institutes of Health, USA) and normalized to anti-PSTAIRE.

2.5.4 Clustering analysis

To differentiate between the different gene expression behaviors during the checkpoint, the high-confidence wild-type periodic gene list was clustered using affinity propagation[150] with the gene expression dynamics for the DNA replication checkpoint (*cdc8*; Figure 2.11) or for the spindle assembly checkpoint ($P_{GAL1\text{-}10}$-*cse4-353*; Figure 2.12). To identify potential transcriptional regulators of each cluster of genes, we performed a transcription factor over-representation analysis based on the *q*-value method[148] using the documented binding information curated by Yeastract[149]. A transcription factor was called over-

represented in a cluster if its *q*-value was less than or equal to 0.01. The top five transcription factors are listed next to each cluster (Figure 2.11,12).

2.6 Chapter 2 collaborator contributions

Work from these studies was done in collaboration with a number of members of the Haase lab and Harer group. The cell-cycle checkpoint experiments were carried out with the help of Laura A. Simmons-Kovacs. Data analysis was completed with the help of Anastasia Deckard and John Harer.

Chapter 3 Computational methods in identifying periodic genes

Gene expression microarrays have proven to be very useful in identifying genome-wide transcriptional behaviors in a variety of different systems and biological mechanisms. These large-scale data sets are so rich with information that may address many different questions about biological processes in addition to the hypotheses of interest. How, then, do biologists parse out the relevant data for a particular hypothesis? The steps taken from raw microarray fluorescence information to analyzing the gene expression levels requires computational algorithms that are able to deal with these large amounts of data. In this chapter, I will discuss pre- and post-processing steps needed in order to ask the seemingly very simple question with genome-wide gene expression dynamics: What portion of the budding yeast genome is cell-cycle regulated?

3.1 Introduction

Gene expression microarrays have demonstrated that a substantial portion of the budding yeast genome is periodically transcribed in normally-cycling cells ([77, 78, 80, 81], reviewed in [87]). In addition to wild-type populations of cells, recent work by Orlando and colleagues [81], Simmons Kovacs and colleagues [133], and myself (as described in Chapter 2) have shown that dynamics of periodic gene expression are maintained in conditions

that perturb oscillations in cyclin/CDK activity and that arrest cell-cycle progression. What are the steps that need to be taken after you hybridize the labeled cDNA to the microarray chips and before you can begin to analyze the resulting data?

3.1.1 Normalization methods

The first step in processing microarray data is to normalize all relevant chips together in order to eliminate non-biological variations in gene expression across time points and/or conditions. Labeled cDNA destined for each chip may vary in mRNA isolation efficiency, fluorescent labeling, and hybridization, for example. Thus, differences in raw values of gene expression of a single gene on a single chip and across multiple chips will be a convolution of biological differences and technical differences. Many of the methods that have been developed to normalize microarrays centers on the design of the actual microarray; a single microarray chip will measure the mRNA levels of thousands of genes for your system of interest. Specific to Affymetrix oligonucleotide arrays, each spot on the array consists of a 25mer probe that is sequence-specific and a perfect match (PM) for a portion of the coding sequence of a gene in the genome. For each gene, a set of eleven to twenty probes will be spotted on the array, covering the span of the coded sequence. A mismatch (MM) probe accompanies each PM probe, in which one base in the PM 25mer is changed to a different base. The set of probes for each gene are scattered throughout the

chip, to prevent from any hybridization biases. How are all of these probes jointly measured to give a single gene expression level for each gene being studied?

The entire process of normalizing microarrays occurs in three main steps: (1) background adjustment, (2) normalization, and (3) summarization [151]. A number of groups have developed varying methods for each of these steps [151]. The normalization method developed by Li and Wong [152, 153] and their subsequently developed tool called dChip will be studied in further detail in this chapter. More detailed descriptions focused on comparing the performances of many of the prominent methods are reviewed elsewhere [151, 154]. Each method makes assumptions concerning each of the steps that lead to comparable gene expression levels for every gene across microarray chips.

In the method developed by Li and Wong [152, 153], no background adjustment across chips is done. Any issues with differing signal intensities within and across chips are handled in the second step, normalization. In the normalization step, the chip with the median overall fluorescent intensity is selected as a baseline to which all other chips to be normalized. The probes of the baseline chip are then compared pairwise to all other chips to define a set of invariant probes that share similar intensities [155]. The probes not considered invariant are then normalized to fit a curve generated by the invariant set of probes. These normalized intensities result in an overall intensity of the non-baseline chip is altered to match the overall intensity of the baseline chip. This

process is completed for every chip in the set to be normalized together, such that the overall intensity of every chip is very similar to each other and the baseline chip.

Once all chips have been normalized to each other, the set of probes that all refer to the same gene on a single chip must be summarized for a final mRNA measurement for that time point or condition. Li and Wong [152, 153] have developed a model based expression index, which creates a weighted average of the probes in a set. In this approach, those probes within a set that display high variability compared to other probes in the same set are down-weighted when calculating the mRNA levels for that individual chip. Often when summarizing the probe sets, researchers will use the accompanying MM probe to identify the binding specificity. Li and Wong [153] originally developed their model to take the MM probe into account. However, the MM probe, while different from the PM probe, is still specific to only a single type of non-specific labeled cDNA that may bind to the PM probe. Any number of sequences could potentially mis-hybridize. Li and Wong [152] further developed their model to only take into account the PM probes in the summarization step. Once normalization and summarization are completed, transcript levels of genes can be directly compared to each other across time or conditions.

3.1.2 Computational algorithms identifying periodic behaviors

With normalized data, researchers can begin the more exciting task of addressing the biological questions that propelled the experiments designed to measure global gene expression. In studies focused on transcription during the cell cycle, observing temporal gene expression dynamics in a synchronous population of cells is critical to identifying periodic transcripts. While it is known that not all genes are periodically expressed, how do we parse out the genes that are cell-cycle regulated? To address this question, computational algorithms that rank gene expression dynamics based on the feature of periodicity have been developed and implemented [78, 79, 133, 143, 144, 156]. Some of these have been used in previous studies to identify periodic genes in wild-type cells [78, 80, 81], while I have used a combination of algorithms as described in Chapter 2. The overlap between the resulting periodic gene lists is surprisingly low [81], and may be explained by a variety of factors, including experimental design and analytical tools. Here I will discuss the differences and similarities between a number of these algorithms.

While experimental design may play a role in affecting the gene expression dynamics, differences between computational algorithms and significance cutoffs play a large role in the differences observed between periodic gene lists. Inherent to these computational methods is the definition of periodicity. This definition that is built into the algorithm is very important and can

have a profound effect on what genes are identified as periodic. Periodicity can be generally described as an event that occurs at regular intervals. To extend this characterization to cell-cycle oscillations, periodic events occur once per cycle at the same point during cellular division. For instance, budding yeast cells undergo budding just prior to the G1-to-S transition while initiating DNA replication during S-phase. These events are periodic relative to the cell cycle. How can we define cell-cycle regulated transcription? Studies have shown that many genes are expressed in a periodic manner throughout the entire duration of a single cell cycle [78, 80, 81, 133]. Thus, all oscillations in gene expression cannot be defined as transpiring during only one phase of the cell cycle. Rather, each periodic gene is reliably expressed at a particular time based on when its transcription is activated. How one defines periodicity will affect which genes are identified as periodic.

Two features play prominent roles in describing periodicity for each computational method that vary in importance of that definition. The first feature addresses the oscillatory nature of gene expression dynamics. Whether by visual inspection or by some mathematical equation, a necessary component of any of these analyses requires identifying genes that are expressed once per cell cycle across multiple cycles. The second feature addresses the dynamic range of the queried genes. This attribute is much more subjective in nature, as it is unclear what minimum dynamic range is above stochastic noise within a cell. With all

computational algorithms, the measurement of periodicity for each gene is given as a ranked list from most periodic to least periodic with some associated score. Differences in how these two features are defined and weighted can affect the final output of the algorithms. It is important to have an understanding of how you want to define periodic genes and select a computational algorithm that best matches your assumptions.

Three different algorithms have been developed in varying fields of study to measure periodic behaviors that I have used to identify periodic genes: de Lichtenberg [79], Lomb Scargle [143, 144], and JTK_CYCLE [156]. Only de Lichtenberg ranks periodicity of gene expression dynamics based on both a periodicity score and an amplitude (referred to as regulation) score at a given period. The periodicity score is rooted in a Fourier transform. The regulation score measures the peak-to-trough of the gene expression profile. This algorithm is a permutation-based method that assigns scores based on randomized input data. While this method is advantageous in terms of testing the null hypothesis that a transcript's mRNA levels are not periodic at a given period, the resulting score will be vary each time the same dataset is run because the permutations will be different every time. Additionally, randomization of the data is also affected when using different sets of genes as different data points can be used for generating permutations. Thus, comparing the periodic gene lists from two different conditions and/or two different sets of genes is difficult when using de

Lichtenberg as the scores were calculated using different sets of randomized data.

Lomb Scargle and JTK_CYCLE do not take amplitude of gene expression into account when calculating a periodicity score and ranking for each gene expression profile. It is unclear what threshold of amplitude constitutes a significant change in gene expression, thus either a tolerant cutoff or no cutoff using amplitude may be most appropriate, depending on the initial hypotheses and questions. Lomb Scargle uses a Fourier transform, similar to de Lichtenberg, at a range of periods given by the user to calculate a periodicity score [142]. Based on this score and the number of time points, a p-value is calculated. As the number of time points increases, the p-value will become smaller. JTK_CYCLE calculates a periodicity score by comparing the directionality of mRNA levels of two points from an entire gene expression profile and then compares that output to sinusoid curves across different periods [156]. Similarly to Lomb Scargle, a p-value is calculated by taking into account the number of time points queried, and will become smaller as more data is tested. The concept that the amount of data affects the confidence of the periodicity score is logical. However, it suggests that the cutoff determined to distinguish between periodic and non-periodic dynamics is arbitrary. In addition to the different definitions of periodicity built into each algorithm, some of the variability in periodic gene lists previously identified may also be due to the chosen significance threshold

chosen by each group. In this chapter, I will address the differences and similarities between these three different computational algorithms and the analytical tools I have used to choose a meaningful cutoff for identifying cell-cycle regulated genes.

3.2 Normalizing temporal gene expression microarrays using dChip

Many of the common normalization methods used for microarray chips have been translated into packages available through Bioconductor on R [157]. Specifically for Affymetrix microarray chips, the *affy* package has been developed to run the method developed by Li and Wong [152, 153] in addition to MAS5, RMA, and GCRMA [155]. After importing all chips that will be analyzed for a study, a single command will run all three steps of background correction, chip normalization, and probe summarization. To normalize using dChip, the command is `expresso(inputdata, normalize.method= "invariantset", bg.correct=FALSE, pmcorrect.method= "pmonly", summary.method= "liwong")`. This command results in no background correction, normalizing chips to a baseline with an invariant set, and summarizing the probes using the model based expression index taking only the PM probes into account. The components of the command can be changed to normalize chips using other methods.

3.2.1 Gene expression in Cdc20-depleted cells is lower than other conditions

Upon first normalizing chips with mRNA related to cells depleted for Cdc20 (P_{GALL}-*CDC20*; CDK "on") with chips for normally-cycling cells, CDK "off" (*clb1,2,3,4,5,6*) cells, the DNA replication checkpoint (*cdc8*^{ts}), and the spindle assembly checkpoint (P_{GAL1}-*cse4-353*), periodic genes appeared to be more lowly expressed than other conditions (Figure 3.1a-c).

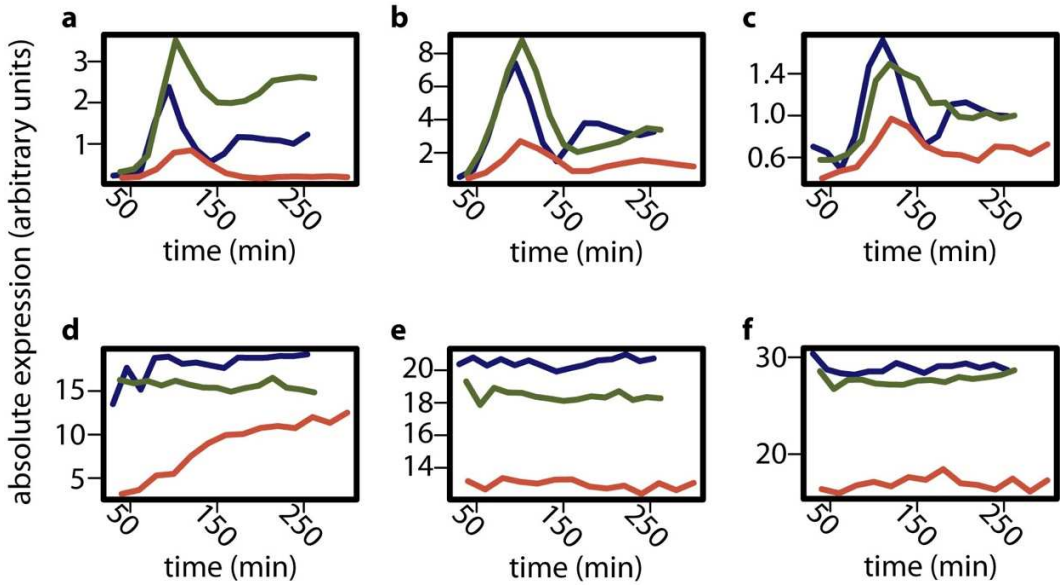

Figure 3.1: Absolute gene expression levels are depressed in CDK "on" cells compared to other experimental conditions. Gene expression curves for wild-type (blue), *clb1,2,3,4,5,6* (green), and P_{GALL}-*CDC20* (red) cells. Periodic genes: *CLN1* (**a**), *POL1* (**b**), and *NDD1* (**c**). Housekeeping genes: *ACT1* (**d**), *CUP5* (**e**), and *CCW12* (**f**). mRNA levels are displays in absolute value, with arbitrary values associated x1000.

Normalization is meant to remove any differences between conditions and time points that are non-biological in nature. Is the lower expression observed in periodic genes a result of biology or an artifact of normalization?

Gene expression dynamics of housekeeping genes should be constitutively expressed during the cell cycle and should be maintained at the same levels across all conditions. Comparing several housekeeping genes reveals that gene expression levels are depressed in CDK “on” cells, similar to what is observed for periodic genes (Figure 3.1d-f). This suggests that the lower gene expression in “CDK” on cells compared to other experimental conditions may be due to some artifact from the normalization process.

To further address the potential for depressed gene expression globally in CDK “on” cells, I compared the mean expression for each gene pairwise between different experimental conditions. This analysis can be further studied by fitting the data points to a linear model to calculate an r^2 value that falls between zero and one. Values close to one indicate that the conditions are similar; whereas values close to zero indicate that the two conditions are not similar. Gene expression levels appear to be very similar in normally-cycling and CDK “off” cells (Figure 3.1). Consistent with this observation, mean expression of transcripts are very similar with r^2=0.968 (Figure 3.2a). Gene expression appears lower in CDK “on” cells compared to both normally-cycling and CDK “off” cells (Figure 3.1). Also consistent with the gene expression dynamics, genes tend to be more highly expressed genome-wide in normally-cycling cells (Figure 3.2b) and CDK “off” cells (Figure 3.2c). Although the r^2 values are not significantly smaller than comparing normally-cycling cells and CDK “off” cells, the mean

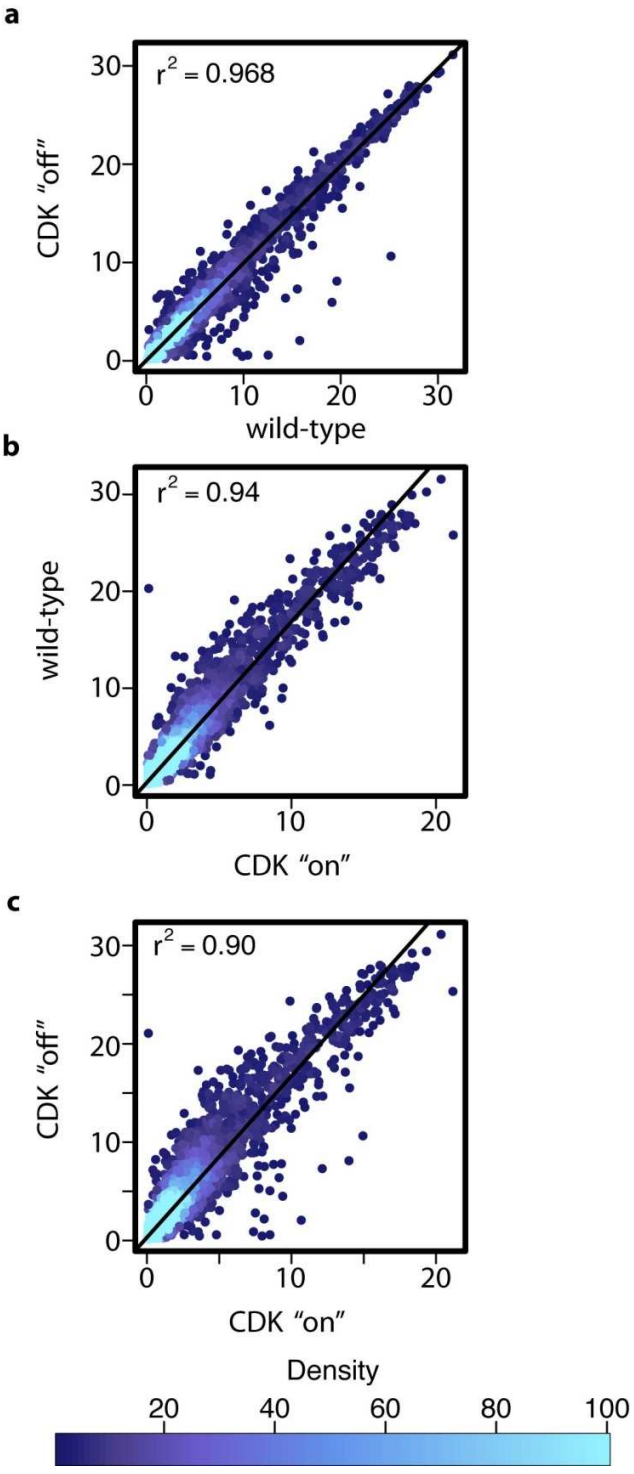
a
r^2 = 0.968
CDK "off"
wild-type
b
r^2 = 0.94
wild-type
CDK "on"
c
r^2 = 0.90
CDK "off"
CDK "on"
Density
20
40
60
80
100

Figure 3.2: Global gene expression is lowered in CDK "on" cells compared to other experimental conditions. Plots comparing the mean expression of each gene between two experimental conditions: normally-cycling cells compared to CDK "off" (*clb1,2,3,4,5,6*) (**a**), CDK "on" (P_{GALL}-*CDC20*) versus normally-cycling cells (**b**), and CDK "on" (P_{GALL}-*CDC20*) versus CDK "off" (*clb1,2,3,4,5,6*) (**c**). A linear model is fit to the data and an r^2 value is calculated and displayed in the plot along with the best fit line.

expression of many genes is skewed to these conditions as seen by more data points above the best fit line (Figure 3.2).

These observations suggest that gene expression is *globally* depressed in the CDK "on" experimental condition. If the lowered gene expression levels were specific to a set of genes, such as only periodic genes, this observation may suggest that some biological factor causes the differences in gene expression measurements. However, the observed global differences suggest an issue with the normalization process.

As described previously, the normalization method developed by Li and Wong [152, 153] and implemented in R [155] consists of two steps: normalization of the chips followed by summarization of the probe sets. To normalize the chips together, the chip with the median overall intensity is selected as a baseline and an invariant set of probes are chosen to correct for differences in intensity between chips. Either or both of these steps could be compromised, which may lead to what appears to be lowered gene expression levels in CDK "on" cells.

3.2.2 Incorrect choice of baseline chip affects resulting global gene expression measurements

Regardless of the order in which chips are entered into the normalization program in R, the baseline chip should always be calculated by measuring the overall fluorescent intensity of each chip. The chips should be ranked based on these measurements and the baseline chip should always be the same. To test this assumption, I took the same chips to be normalized together, but shuffled the order in which I input the chips. Table 3.1 describes the orders of the chips input into the normalization algorithm.

Table 3.1 Orders of chips input for normalizations

Input order	Normalization 1	Normalization 2
1	Wild-type replicate 1	Wild-type replicate 1
2	Wild-type replicate 2	Wild-type replicate 2
3	CDK “off” replicate 1	CDK “on” replicate 1
4	CDK “off” replicate 2	CDK “on” replicate 2
5	CDK “on” replicate 1	CDK “off” replicate 1
6	CDK “on” replicate 2	CDK “off” replicate 2

The true chip with the median overall intensity is from the second wild-type replicate that contains measurements for cells 126 minutes after release into the experiment. However, when comparing the two orders of chips used for

normalization, this chip was not identified for either of the orders. Rather, for normalization order 1 (Table 3.1), the chip for the second replicate of CDK “on” cells 40 minutes after release into the experiment was labeled the median intensity chip. A different baseline chip was also selected for the second order tested (Table 3.1): the chip for the second replicate of CDK “off” cells 30 minutes after initiating the experiment. Why is the incorrect baseline chip being selected in both orders used to normalize the data together?

In both cases, the only element differing between the two normalizations is the order of the chips. Further inspection of the orders demonstrates that the chosen baseline chips for each respective order is the chip that halfway through the list. This finding suggests that the code in R implemented to rank the chips based on overall intensity contains a bug of some kind. In fact, a bug in the code fails to reorder the chips based on overall fluorescent intensity and uses the order input by the user as the ranked list to choose the baseline chip. Altering the code to reflect the correct ranking of chips results in the correct and true baseline chip for both orders listed in Table 3.1. Since discovering this bug, it has since been fixed in the affy package and will choose the correct baseline chip regardless of the order in which you input your arrays to be normalized.

3.2.3 Masking *Schizosaccharomyces pombe* probes is best practice for characterizing an invariant set of probes for normalization

Another aspect of normalization not taken into account in the preprocessing steps coded in the affy package through Bioconductor on R is the makeup of the Affymetrix Yeast 2.0 oligonucleotide arrays. The microarrays used in the experiments described in Chapter 2, and previously [81, 133], used these chips and do not contain probes only specific to *Saccharomyces cerevisiae*. Probes specific for *Schizosaccharomyces pombe* can also be found on these chips. Thus, there may be some, but likely very little, cross-hybridization of mRNA isolated from budding yeast to probes meant for fission yeast. Most cross-hybridization likely stems from genes that are strongly conserved across these two species, including housekeeping genes.

How could the fission yeast-specific probes affect normalization of mRNA measurements for budding yeast experiments? During the chip normalization process, a set of invariant probes is chosen that are similar in spot intensity in both the baseline chip and the chip being normalized. It could be assumed that the amount of cross-hybridization to *S. pombe* probes is uniform across all chips being normalized. This may lead to a large number of probes specific for fission yeast in the invariant set for the normalization step. This result may affect the outcome of the actual normalizations and the subsequent summarization steps.

To test this hypothesis, I normalized chips from the experimental conditions and the order listed in Table 3.1 for order 1 either with the *S. pombe* probes removed before normalization or without removing the *S. pombe* probes. If the fission yeast probes do not have an effect on the outcome of the normalization, then a comparison of the two methods by fitting to a linear model should yield an r^2 value of one. However, if the probes meant for fission yeast does affect the normalization outcome, an r^2 value less than one would be calculated. I find that when comparing the same dataset with and without fission yeast probes, the r^2 value is very close to one, but not quite (Figure 3.3). This is the case for all experimental conditions normalized: normally-cycling cells (Figure 3.3a), CDK "off" cells (Figure 3.3b), and CDK "on" cells (Figure 3.3c). The resulting fits from this analysis suggest that while probes specific for *S. pombe* do affect the resulting gene expression levels, the effect is very small. Thus, while removing fission yeast probes may not be necessary, it is best practice for accurate mRNA measurements for experiments done in budding yeast.

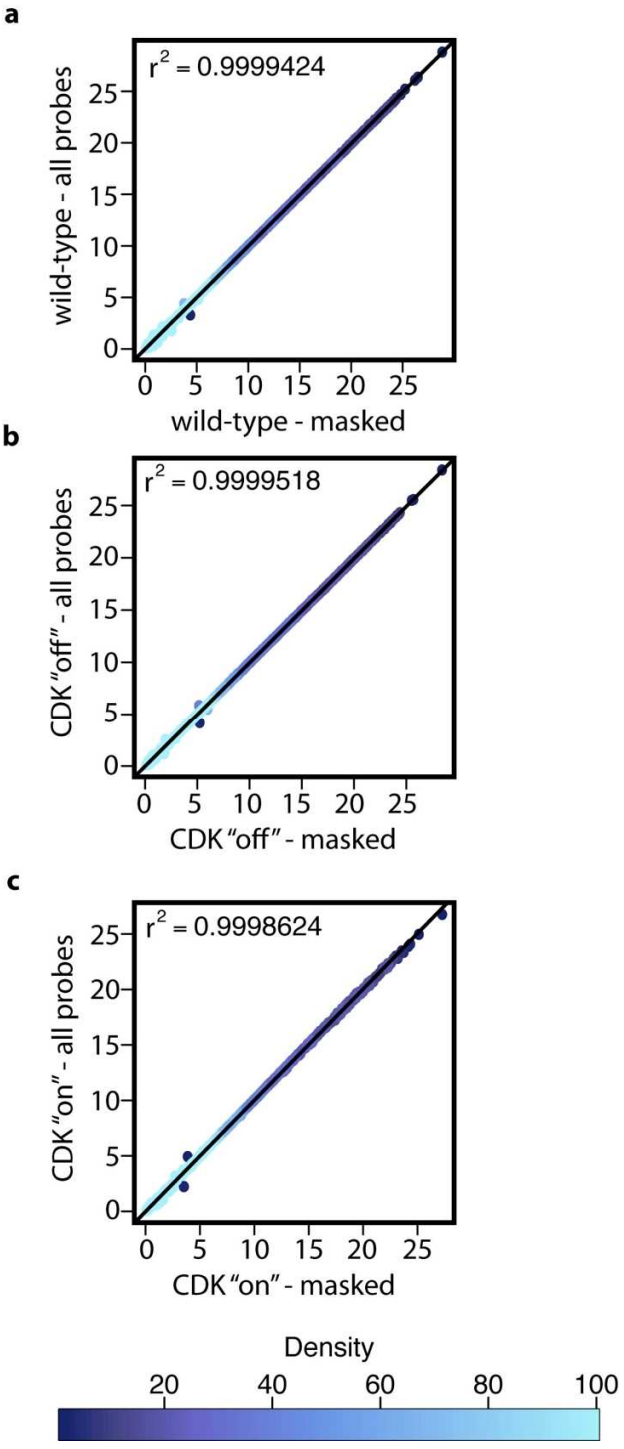
a
r² = 0.9999424
wild-type - all probes
wild-type - masked
b
r² = 0.9999518
CDK "off" - all probes
CDK "off" - masked
c
r² = 0.9998624
CDK "on" - all probes
CDK "on" - masked
Density
20
40
60
80
100

Figure 3.3: Removing probes specific to *Schizzosaccharomyces pombe* affects the resulting gene expression measurements during normalization. Plots comparing the mean expression of each gene between two normalization approaches: with and without fission yeast-specific probes in normally-cycling cells (**a**), CDK "off" cells (*clb1,2,3,4,5,6*) (**b**), and CDK "on" cells (P_{GALL}-*CDC20*) (**c**). A linear model is fit to the data and an r^2 value is calculated and displayed in the plot along with the best fit line.

3.2.4 Different baseline chips do not affect the log-fold change in expression, only absolute values

A major advantage to the method developed by Li and Wong [152, 153] and the implementation in the affy package [155] is the ability to calculate absolute values with some associated arbitrary unit. While it does not reflect the number of individual mRNAs of each expressed gene, it does allow for comparisons of levels of genes across time within and across experimental conditions. The issues identified in the above sections suggests that differences in gene expression levels at the absolute level may not have been due to biological mechanisms, but rather due to the normalization algorithm itself.

Another way to visualize data is to look at fold changes of the gene expression dynamics of a transcript and compare it to other transcripts in the same experimental condition or to the same transcript across multiple experimental conditions. Do the normalization-specific differences affect both the absolute value and fold change of gene expression dynamics? To address this question, the gene expression dynamics of the genes in Figure 3.1 were converted to $\log_2$(expression/mean) for each gene and experimental condition. I

find that although gene expression levels appear depressed in CDK “on” cells compared to both normally-cycling and CDK “off” cells (Figure 3.1), the fold changes are very similar between all three conditions (Figure 3.4).

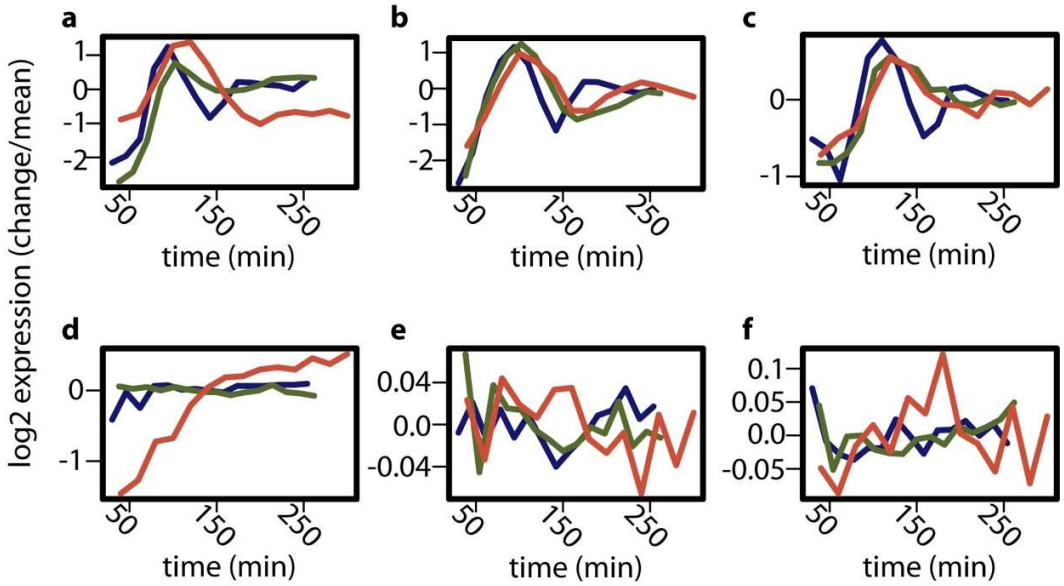

Figure 3.4: Fold changes in gene expression levels are the same across experimental conditions. Gene expression curves for wild-type (blue), *clb1,2,3,4,5,6* (green), and P_{GALL}-*CDC20* (red) cells. Periodic genes: *CLN1* (**a**), *POL1* (**b**), and *NDD1* (**c**). Housekeeping genes: *ACT1* (**d**), *CUP5* (**e**), and *CCW12* (**f**). mRNA levels are displays in $\log_2$fold dynamics.

Further, comparing two different normalizations – with or without the correct baseline chip selected – shows that although the absolute values vary, the fold changes remain the same in both normalization processes (Figure 3.5).

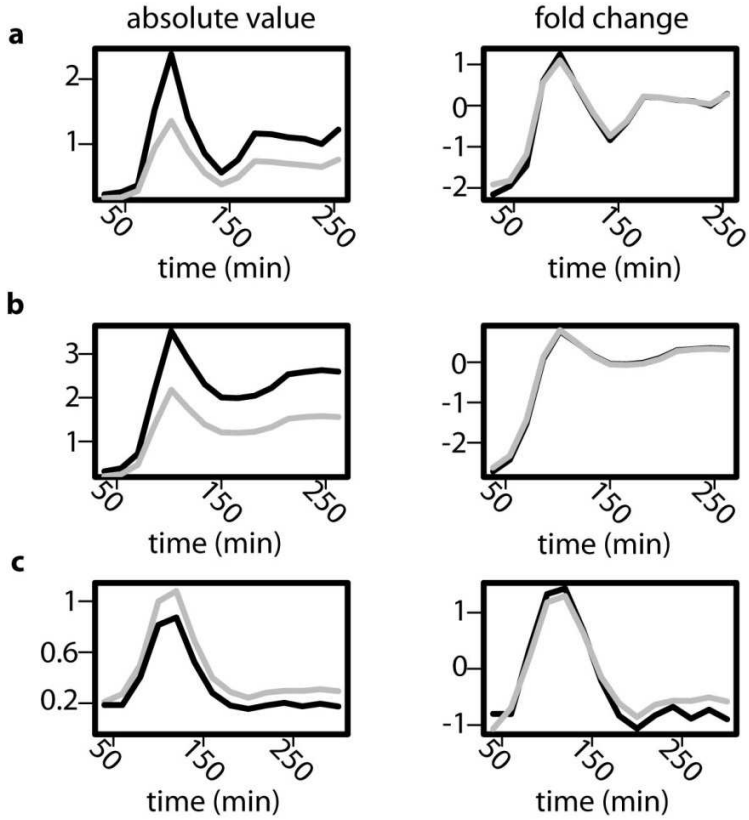

Figure 3.5: Fold changes in gene expression levels are the same across varying normalization parameters. Gene expression curves for the transcript *CLN1* in wild-type (**a**), *clb1,2,3,4,5,6* (**b**), and P_{GALL}-*CDC20* (**c**) cells. Black, incorrect baseline chip; gray, correct baseline chip. Left column displays mRNA levels in absolute values with arbitrary values associated x1000. Right column displays mRNA levels in $\log_2$fold dynamics.

These results suggest that the problems identified the normalization process mostly affects the absolute gene expression calculations. Thus, if normalization is not done properly, incorrect conclusions can be made about differences in gene expression levels across time and experimental conditions. However, findings based solely based on the behavior of the gene regardless of relative mRNA levels will not change based on the normalization methods chosen.

3.3 Identifying cell-cycle regulated genes using periodic-detecting algorithms

Regardless of the many studies to classify cell-cycle regulated transcription in budding yeast [77, 78, 80, 81], the question of what portion of the genome is periodic remains a contentious question in the field. Comparing three periodic gene lists [78, 80, 81], Orlando and colleagues only found 440 genes shared among all studies. Much of the discrepancy observed between the different studies relates to the computational approaches to identifying periodic genes. In this section, I will go into detail describing some of the issues to take into consideration when identifying periodic genes using three exemplar algorithms – de Lichtenberg [79], Lomb Scargle [143, 144], and JTK_CYCLE [156]. I will also outline analytical tools that can be used in addressing each issue.

3.3.1 Selecting a period to query for periodic transcripts affects the resulting gene list

Inherent in studying periodic behaviors is the idea of searching for a set of genes that oscillate at a particular period. Specific for the cell cycle, periodic genes are expressed only once per cellular division. How do we calculate the cell-cycle period in normally-cycling cells? In the Haase lab, we use a statistical model referred to as CLOCCS that takes takes budding indices from synchrony-release experiments of wild-type cells and, with a set of assumptions, will infer a

number of parameters [158]. These parameters include mother cell-cycle period, the amount of extra time daughter cells remain in G1, and the amount of synchrony of the population [158]. These learned parameters can then be used to query periodic genes at specific periods.

Orlando and colleagues used the mother cell-cycle period to identify cell-cycle regulated genes using the de Lichtenberg algorithm [79] (Table 3.2).

Table 3.2: Summary of CLOCCS parameters for wild-type cells and the calculated average cell-cycle period

Wild-type Replicate	Mother Period	Daughter G1	Average Period
1	77.1	41.4	97.8
2	85	35.1	102.6

Using a p-value cutoff of $p \leq 0.2$, 1275 periodic genes were identified [81]. The population of cells used for the experiment was synchronized to early G1 using centrifugal elutriation. In wild-type cycling cells, the population is a heterogeneous mix of both mother and daughter cells. In a synchronous population of cells, the first cycle may be close to the mother cell-cycle period but will be a convolution of two different periods after one round of division. To address this potential issue, I calculated an average cell-cycle period using the following equation: (mother period) + (mother period+daughter-specific G1)/2.

This results in average cell-cycle periods that are longer than the mother period used to measure periodicity (Table 3.2). Using the de Lichtenberg

algorithm again and a p-value cutoff of $p \leq 0.2$, I identified 1912 cell-cycle regulated genes at the longer, average period.

What are the differences between the two periodic gene lists generated by de Lichtenberg at the two periods queried? After comparing the genes that fell below the threshold, 1010 genes are classified as periodic in both resulting lists (Figure 3.6a,c). This result suggests that periodic genes may be scored as periodic over a range of periods. The average cell-cycle period adds nearly 1,000 periodic genes in addition to the shared periodic genes. These genes appear periodic and are expressed at all phases of the cell cycle (Figure 3.6b). Only a small set of genes were identified as periodic only by the mother cell-cycle period (Figure 3.6d). Using the average cell-cycle period has yielded the largest periodic gene list in any study, suggesting that up to 30% of the yeast genome may be periodically expressed in cycling cells.

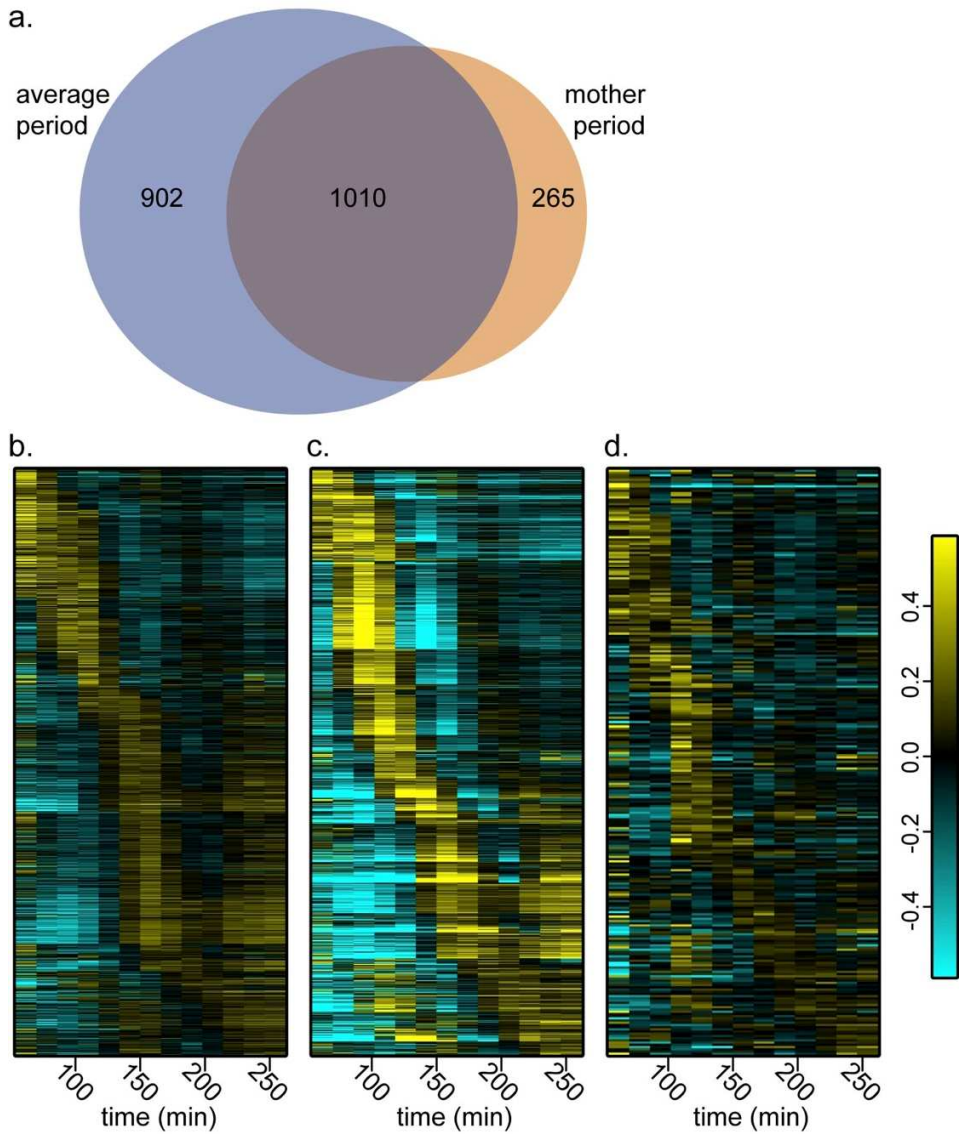

Figure 3.6: Comparing periodic gene lists from de Lichtenberg using two different periods. Venn diagram depicting the overlap between the periodic gene lists that were generated by de Lichtenberg at the mother cell-cycle period or the average cell-cycle period for wild-type cells (**a**). Heat maps showing the mRNA levels of genes that are periodic only at the average cell-cycle period (**b**), that are periodic at both periods (**c**), and that are periodic only at the mother cell-cycle period (**d**). Each heat map displays a different set of genes. Transcript levels are depicted as $\log_2$-fold change relative to the mean expression.

How do varied periods affect the output of other algorithms? Lomb Scargle and JTK_CYCLE do not require a single period input to measure periodicity, but rather query a range of periods given by the user [142, 156]. For both algorithms, periods spanning from 50 minutes to 200 minutes were queried. Testing the effect of varying periods can be easily done as a p-value for each period in the range is calculated. At a number of different p-value cutoffs, the peak number of genes classified as periodic falls at the average cell-cycle period and falls off with some normal distribution (Figure 3.7).

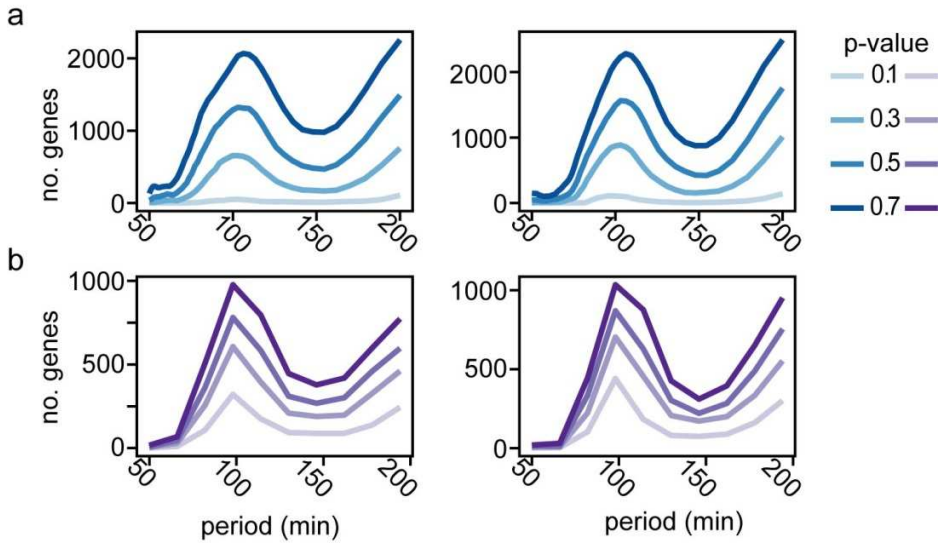

Figure 3.7: Size of gene list varies across periods and p-value cutoffs for periodic-detecting algorithms. Histograms depicting the number of genes included in wild-type periodic gene lists in replicate 1 (left) and replicate 2 (right) for Lomb Scargle (**a**) and JTK_CYCLE (**b**).

The number of genes begins to increase again as the large periods are queried, which is due to the queried period being longer than the experimental time

course. How do the calculated p-values change over the course of the period range? At the periods close to the average cell-cycle period, p-values associated to periodic genes become small and are larger as the queried period increases in distance from the average period (Figure 3.8a,b).

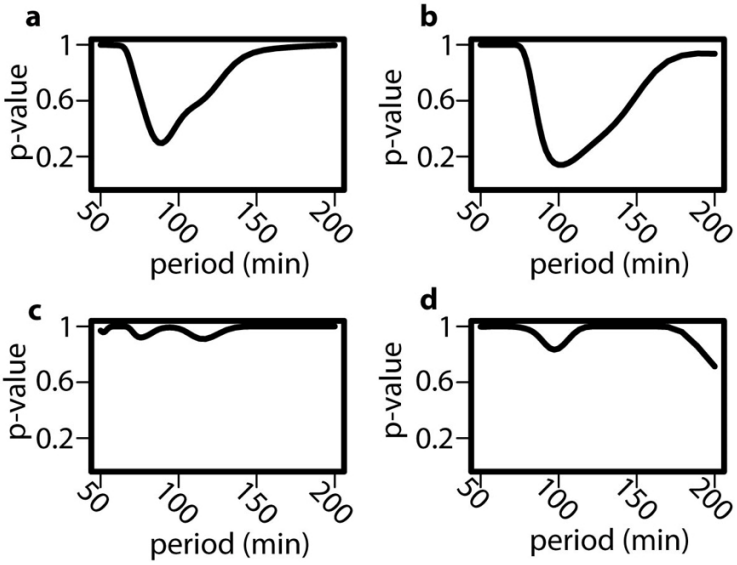

Figure 3.8: Periodograms of periodic and non-periodic genes from Lomb Scargle. Plots of calculated p-values across a range of periods queried for wild-type replicate 1 by Lomb Scargle. Periodic genes: *CLN1* (**a**), *CLB2* (**b**). Non-periodic genes: *ACT1* (**c**), *CCW12* (**d**).

However, genes that are not periodic do not demonstrate this behavior (Figure 3.8c,d). This result suggests that within some period range and p-value cutoff, the resulting periodic gene lists will not vary much.

3.3.2 Different factors can affect periodicity scores and p-values

After the periodicity scores are calculated, these scores can then be transformed into p-values. For all three algorithms, different factors can be varied

that will change the resulting p-values. de Lichtenberg calculates an amplitude score, referred to as a regulation score, in addition to the periodicity score. Further, this algorithm can take multiple experiments into account when calculating the final score. The final score de Lichtenberg calculates can change depending on the weight given to the two scores: periodicity and regulation. If periodicity is more important, it is weighted such that even if a gene does not show high amplitude, it may still receive a good overall score. The opposite is true if the regulation score is weighted more heavily. Thus, depending on the importance given to each measurement, the final ranking of genes and associated scores will vary.

As mentioned previously, Lomb Scargle and JTK_CYCLE do not take amplitude into account when calculating a final p-value. Rather, the p-value is calculated by taking into account the periodicity score and the number of time points queried [142, 156]. The logic here follows that as more time points are included and the gene expression profile receives a high periodicity score, the more confident the algorithm is in its periodicity score. Comparing resulting p-values from normally-cycling cells with 13 time points or fit to a cubic spline to increase the number of time points to 50 time points shows that the dynamic range of both the periodicity scores and the resulting p-values are very different (Figure 3.9). This suggests that choosing a significant p-value cutoff for Lomb Scargle and JTK_CYCLE may not need to fall below traditional cutoffs.

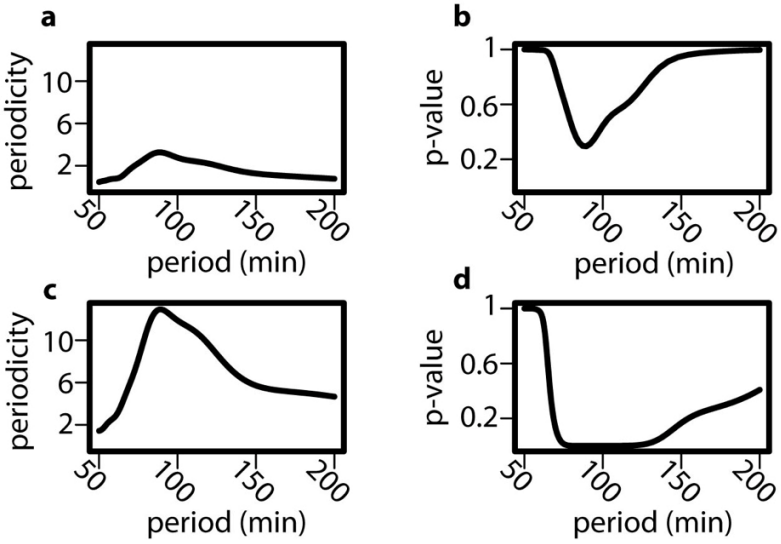

Figure 3.9: Splining data for more time points affects both the periodicity score and the p-value. Plots of periodicity scores across a range of periods queried for wild-type replicate 1 by Lomb Scargle for non-splined data (**a**) and splined data with 50 time points (**c**). Plots of calculated p-values across a range of periods queried for wild-type replicate 1 by Lomb Scargle for non-splined data (**b**) and splined data (**d**).

3.3.3 Choosing a threshold to establish a set of cell-cycle regulated genes

The output of a period-detecting algorithm is a list of ranked genes from most periodic to least periodic that have some associated score. There is no one threshold to use to delineate between periodic and non-periodic behaviors. Whatever cutoff is chosen, non-periodic genes may lie above the cutoff while periodic genes lie below the cutoff. Additionally, the algorithms do not relate that information. It is up to the researcher to choose some cutoff that best represents the global behavior queried in the experimental condition and the algorithmic parameters. In the above section, it was described that the p-values can change

depending on weighting parameters or number of time points. Thus, even choosing some ‘significant’ p-value can also be an arbitrary manner in which thresholds are selected. Given the amount of uncertainty, is there a way to choose a threshold and period range that will best represent the biology?

One analytical tool to help in selecting an appropriate p-value cutoff and period range for identifying cell-cycle regulated genes is to compare the inclusion of a test set of known periodic genes in gene lists generated using different p-value cutoffs and period ranges. When developing their permutation-based model, de Lichtenberg and colleages [79] built a list of known periodic genes to determine how well the algorithm performed. This same type of analysis can be extended to choosing a set of cell-cycle regulated genes. To compare to the de Lichtenberg test gene list, gene lists were created from varying p-value cutoffs and period ranges from both Lomb Scargle and JTK_CYCLE. Comparing these to the de Lichtenberg test set shows that more of the test genes are included in the periodic gene lists as both the p-value cutoffs become more lax and the period range is expanded (Figure 3.10a,b). In the field of cell-cycle regulated transcription, an outstanding question is learning the transcription factors (TFs) that play a role in regulating periodic transcription. Are the TFs known to affect cell-cycle regulated transcription included in these gene lists created? To address this question, an additional test set was constructed to include many cell-cycle TFs in addition to other non-transcriptional cell-cycle regulators. Similar

to the de Lichtenberg test set, the TF test set had more genes included as the p-value cutoffs and the period ranges grew larger (Figure 3.10c,d). This analysis suggests that choosing a threshold at which to delineate between periodic and non-periodic is an arbitrary task that can be aided with previous biological knowledge.

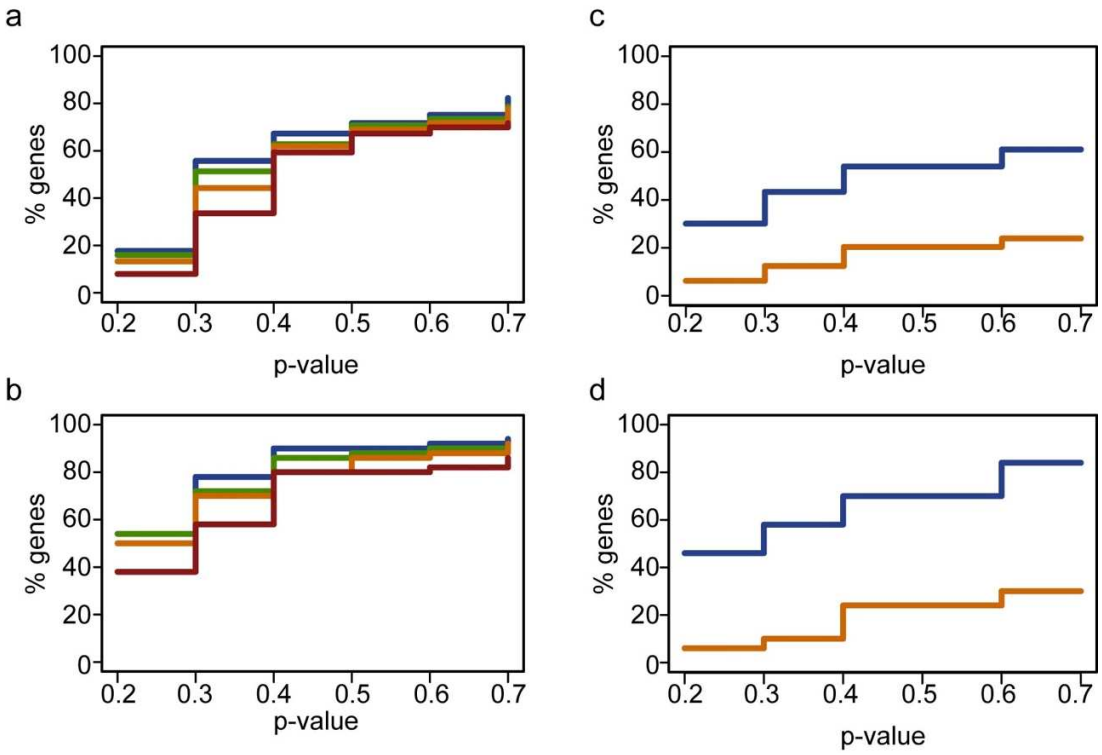

Figure 3.10: Comparing test periodic sets to gene lists generated by Lomb-Scargle and JTK_CYCLE. Number of genes from deLichtenberg test set (a) or transcription factor test set (b) found in Lomb-Scarge periodic gene lists varying in p-value cutoff and period range. Blue, 87.2-106.3 minutes; green, 91.9-103 minutes; orange, 94.4-100 minutes; red, 97 minutes. Number of genes from deLichtenberg test set (c) or transcription factor test set (d) found in JTK_CYCLE periodic gene lists varying in p-value cutoff and period range. Blue, 82-114 minutes; orange, 98 minutes.

3.3.4 Different periodic algorithms score periodicity differently

As discussed previously, a major effort to understand cell-cycle regulated transcription entails identifying the portion of the genome that is periodically expressed. Previous studies have attempted to define the genes that oscillate during the cell cycle, but comparing them to each other shows that there is very little overlap with each other. This outcome is mostly due to the differences in the definition of periodicity. Each group used different criteria to characterize the elements of a periodic gene. What are the specific differences between the resulting periodic gene lists from varying algorithms and what are the assumptions built into the algorithms that produces those disparities?

After choosing a cutoff and period range, how similar are the periodic gene lists generated by de Lichtenberg, Lomb Scargle, and JTK_CYCLE? Table 3.3 lists the thresholds chosen for each gene list and the period ranges that were chosen with the help of the test periodic gene lists (Figure 3.10).

Table 3.3: Chosen period ranges and p-value cutoffs for de Lichtenberg, Lomb Scargle, and JTK_CYCLE.

Algorithm	Period Range	p-value cutoff
de Lichtenberg	Average period	0.2
Lomb Scargle	92-103 min	0.5
JTK_CYCLE	98 min	0.5

In total, 458 genes are shared between all three algorithms (Figure 3.11). While this represents approximately 50% of the genes, this leaves a number of genes that are specific to one or two of the algorithms.

To understand the differences between the algorithms, further studying the non-overlapping periodic genes provide the best information. Lomb Scargle and JTK_CYCLE show a large amount of overlap (Figure 3.11a), so these two algorithms will be compared together as a unit to de Lichtenberg. The de Lichtenberg algorithm classifies periodic genes that are much higher in amplitude compared to Lomb Scargle and JTK_CYCLE (Figure 3.11b,d). The lower amplitude observed in Lomb Scargle and JTK_CYCLE is due to no weight on amplitude when calculating periodicity, while de Lichtenberg does take amplitude into account. Another result of the amplitude score for the de Lichtenberg-only periodic gene list is a set of genes that peak once at the beginning of the time course and is not expressed subsequently (Figure 3.12a). Even if the amplitude score is down-weighted in importance compared to the periodicity score, genes with very large amplitude changes will overwhelm the scoring and fall within the threshold for choosing a periodic gene list. These genes are excluded from Lomb Scargle and JTK_CYCLE outputs. Another set of genes included in the de Lichtenberg periodic gene list and excluded from the other two algorithms are genes expressed in late M phase to early G1. These genes are very periodic, however, they are likely scored poorly as less than 1.5 cycles are measured

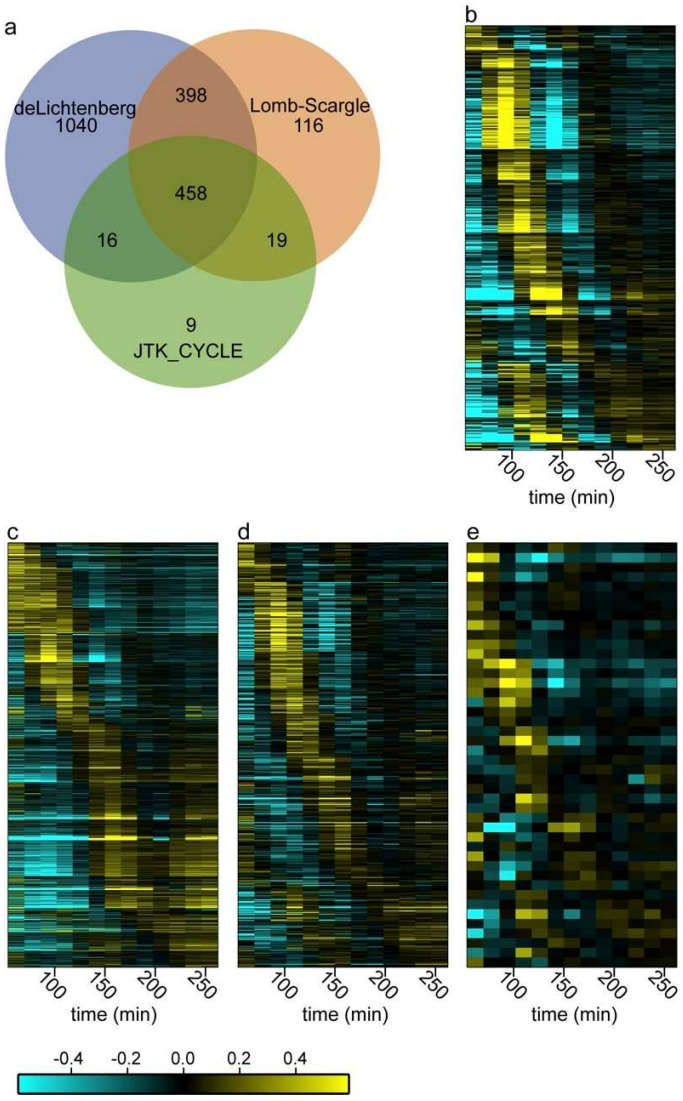

Figure 3.11: Consensus of deLichtenberg, Lomb-Scargle, and JTK_CYCLE. Venn diagram of the resulting periodic gene lists from deLichtenberg (blue), Lomb-Scargle (orange), and JTK_CYCLE (green) (a). Heat maps showing the mRNA levels of genes that are periodic in all gene lists (b), that are periodic in deLichtenberg and do not overlap with all gene lists (c), that are periodic in Lomb-Scargle and do not overlap with all gene lists (d), and that are periodic in JTK_CYCLE and do not overlap with all gene lists (e). Each heat map displays a different set of genes. Transcript levels are depicted as $\log_2$-fold change relative to the mean expression.

(Figure 3.12b). Upon further inspection of the scoring outputs for this set of genes in all three algorithms, Lomb Scarge and JTK_CYCLE do not assign favorable p-values (Figure 3.12c,d). However, de Lichtenberg ranks these genes' periodicity score very low and their amplitude score very high (Figure 3.12e). This result suggests that regardless of the algorithm, M/G1 genes do not receive good periodicity scores due to the lack of data.

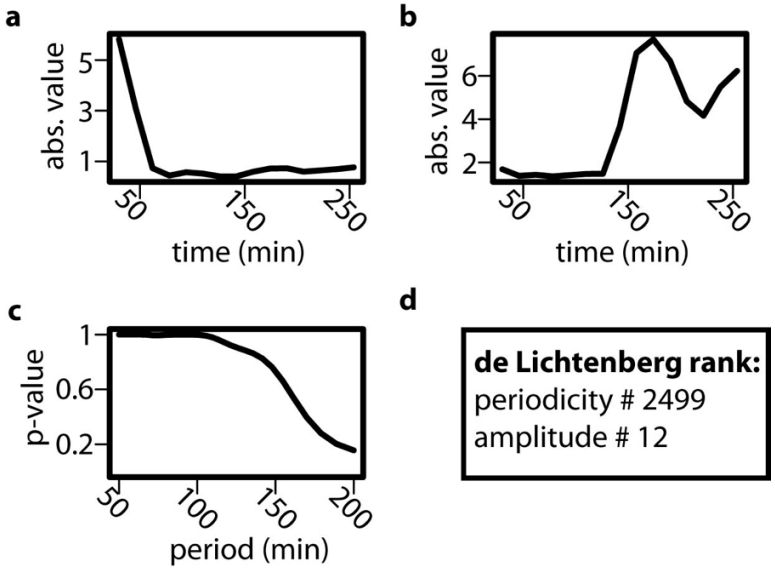

Figure 3.12: Differences between de Lichtenberg, Lomb Scargle, and JTK_CYCLE outputs. Gene expression curves in wild-type cells for *LAP4* (**a**) and *CTS1* (**b**). mRNA levels are displays in absolute value, with arbitrary values associated x1000. Plot showing the p-values from Lomb Scargle associated with *CTS1* gene expression dynamics (**c**). Ranking of *CTS1* periodicity and amplitude scores among all budding yeast genes calculated by de Lichtenberg for wild-type replicate 1 (**d**).

This analysis shows that no algorithm is perfect and using some combination of these, and other, algorithms may allow for a more comprehensive

or restrictive periodic gene list. The best choice of algorithm is highly dependent on the biological question being asked and the approach being taken.

3.4 Discussion

While normalization and identifying cell-cycle regulated transcripts seems like a very straightforward task, there are many details that, if overlooked, can result in improper interpretation of data. In this chapter, I detail the work done to attempt to better understand the normalization processes and the algorithms used to rank genes based on periodicity.

3.4.1 Normalization is an important preprocessing step in data analysis

Many normalization processes have been proposed and studied for normalizing Affymetrix oligonucleotide arrays. While each is different from each other, the principle is the same: to remove any non-biological differences. The goal of normalization is to make mRNA measurements comparable across time points and experimental conditions. Normalization methods have been converted into code that is implemented through packages on different coding platforms [155]. While this is very helpful for biologists who may not understand how to implement an algorithm, the danger is turning normalization into a black box that takes your data and spits out results that are trusted immediately for interpreting biological mechanisms. In this chapter, I describe just one issue that was discovered by wondering why only one condition showed lower global mRNA

levels compared to other experimental conditions. It is important to have a working knowledge of the algorithms chosen for both pre- and post-processing steps.

3.4.2 Defining cell-cycle regulated transcription

While great efforts have been made to define cell-cycle regulated transcription, very little overlap exists between the different resulting gene lists. With careful analysis, I have been able to detail how and why three different algorithms produce such different results. Not only can periodic gene lists be affected by the queried period (Figures 3.6, 3.7, and 3.8), but also by the threshold chosen (Figures 3.10, 3.11, and 3.12). Perhaps different threshold cutoffs would make different gene lists more similar. However, the takeaway from these analyses is that there is not *the* periodic gene list, but several different periodic gene lists that depend on biological question, algorithm choice(s), and selected thresholds.

3.5 Future Directions

In this chapter, I discuss the differences between different periodic-detecting algorithms. Future work must be continued to further develop and understand the methods used to identify periodic genes.

3.5.1 Are p-values the most efficient way to choose cutoffs?

While calculating p-values is a very useful tool to measure the significance of differences between conditions and other biological measurements, are p-values the best mechanism to calculate a cutoff for distinguishing between periodic and non-periodic behaviors? As discussed in section 3.3.2, p-values can be calculated differently when parameters are varied. The connotation of some 'significant' p-value loses meaning when the parameters can be changed just by different weighting or by splining data. The user makes these changes and thus different users will have different outputs depending on stringency of parameters. For each of the algorithms, is there a better way to define periodic transcription? Further work and analyses must be done to eliminate the need for p-values or to streamline the ways we choose p-values without allowing for so much variability.

3.5.2 How are biological replicates used to better inform periodic gene lists?

Of the three periodicity-detecting algorithms, de Lichtenberg is the only one capable of integrating scores for multiple biological replicates. Replicates are a very useful tool in any experiment, lending confirmation, or not, of observations found in one or the other replicate. Currently, biological replicates are run through Lomb Scargle and JTK_CYCLE separately with the exact same parameters. How the outputs from the replicates are integrated has not yet been established. One method that I have implemented is taking only genes that fall within the threshold

parameters in both replicates. This ensures that genes identified from either algorithm are more likely to be truly periodic. Rather than simply taking an intersection of gene lists, is there a more principled method to integrate the scores from each biological replicate? Perhaps using the same principles used in de Lichtenberg will provide some insight into some practices that best utilizes all data available to address cell-cycle regulated transcription.

3.6 Experimental Details

3.6.1 Normalization methods

CEL files from the Affymetrix Yeast 2.0 oligonucleotide arrays generated for wild-type, CDK "off" (*clb1,2,3,4,5,6*) and CDK "on" (*cdc20Δ;*P_{GALL}*-CDC20*) were normalized and summarized using a modified version of the dChip[148] method from the `affy` package (v. 1.32.1) in Bioconductor [155, 157] within the R statistical programming environment. The probes specific to *S. pombe* were removed using the s_cerevisiae.msk file from Affymetrix before normalizing the CEL files.

3.7 Chapter 3 collaborator contributions

Much of the analysis done in this chapter were completed in collaboration with other members of the Haase lab and collaborating groups at Duke University. The analysis done to identify the bug in the dChip code in the *affy* package was completed with Michael Mayhew. Analysis done to compare

different periodicity-detecting algorithms was done in close collaboration with Anastasia Deckard and John Harer.

Chapter 4 Discussion and Future Directions

In this dissertation, I have addressed questions focused on better understanding how a transcription factor (TF) network and the periodic transcriptional program are coupled to other events with the proper timing during the cell cycle. Two alternative models have been proposed to explain how a TF network is regulated to couple the timing of its oscillations with cell-cycle events in wild-type cells. The first model proposes that oscillations in cyclin/CDK activity entrain a TF network, in addition to other events, to maintain a proper period (Figure 1.1b) [132] (reviewed in [134]). The second model suggests that a TF network serves as the underlying cell-cycle oscillator that is responsible for the timing of cell-cycle events by regulating the timing of cyclin synthesis and other periodic genes (Figure 1.1c) [133]. Both of these models were built around the same data, yet two very different models could explain previous findings. The data presented in this dissertation distinguishes between these two models and suggests that the second model fits all current observations. Further, we have added an extra layer of control with the discovery that cell-cycle checkpoints have the ability to arrest transcriptional oscillations. In this chapter, I will discuss the significance of these findings, and the future work needed to further elucidate the regulatory modules that control the TF network oscillator.

4.1 A transcription factor network does not need to be regulated by CDK activity

Previous work has shown that in the absence of cyclin/CDK activity, transcriptional oscillations continue [81, 133]. A transcription factor network was proposed to control periodic transcription and to act as the cell-cycle oscillator that keeps the time of events [81, 133]. Further experiments have shown that perturbations to the TF network via over-expression or deletion of network TFs can alter the period of cell-cycle oscillations in the absence of S-phase and mitotic cyclins [81, 133]. This finding suggests that a TF network is, in fact, a part of the underlying cell-cycle oscillator. These data fit a model in which at TF network acts as an oscillator and controls temporal order via periodic transcription, including cyclins (Figure 1.4b) [133]. However, a TF network cannot trigger cell-cycle events; cyclins, when bound to CDK, are capable of triggering these events at the proper time (Figure 1.4b). In addition, cyclin/CDK activity is known to affect the activity of many of the network TFs, resulting in feedback onto the function of the TF network oscillator.

This same set of data has been interpreted in a different way, with a CDK-centric perspective. In addition to CDK-independent transcriptional oscillations, other CDK-autonomous cell-cycle oscillations have been observed, sometimes with non-wild-type periods [28, 81, 131, 132]. How then are all of these autonomous oscillations coupled to each other during a wild-type cell cycle? The

importance of oscillations in cyclin/CDK activity in triggering cell-cycle events is very well studied in all organisms. A recent study proposed that cyclin/CDKs are important for entraining these autonomous cell-cycle oscillations such that they occur not only with the correct period, but also at the correct time [132] (reviewed in [134]. Given this model, the only situation in which these autonomous oscillations can be observed is in the absence oscillations in cyclin/CDK.

In chapter 2, I address these two models and design an experiment to delineate between them. If autonomous transcriptional oscillations can only be revealed in the absence of cyclin/CDK activity, than a substantially different transcriptional response should be observed when cyclin/CDK activity is persistent. However, if the first model is correct, then at least some portion of transcriptional oscillations should continue. In fact, we see that when mitotic cyclin Clb2 is stabilized, transcriptional oscillations can be observed (Figure 2.7). This result fits best with the TF network oscillator model. However, we do know that TFs known to be regulated by Clb2/CDK are affected, as gene expression of their targets lose periodicity (Figure 2.5).

A most striking finding from this experiment is that the period of transcriptional oscillations in CDK “on” cells is exactly the same as those in CDK “off” cells (lacking S-phase and mitotic cyclins). This result suggests that regardless of the state of cyclin/CDK levels, a TF network continues to oscillate at the same period. This data does not completely fit with recent observations; as

more cyclin/CDK activity is removed from cells, the period of transcriptional oscillations lengthens [133]. This suggests that cyclin/CDKs have the ability to fine tune TF network oscillations. How does the absence of S-phase and mitotic cyclin/CDK activity and persistent mitotic cyclin/CDK activity result in the same period of transcriptional oscillations?

One possible explanation for this puzzling observation is that while persistent Clb2/CDK activity does have known effects on the TF network oscillator, the TFs that are not regulated by Clb2/CDK are in a similar cellular state compared to cells lacking S-phase and mitotic cyclin/CDK activity. Regardless of whether cyclin/CDK activity is oscillating, absent, or persistent, only a subset of TFs is affected by this activity. It could be that a sufficient number of TFs that act throughout the cell cycle are capable of generating transcriptional oscillations. This could explain why a similar period is observed in two very opposite conditions.

Another potential explanation for this finding is centered on the concept of positive feedback. Two positive feedback loops between network TFs and cyclin/CDK activity exist. The first positive feedback loop is built between the G1/S TF SBF (SCB Binding Factor) and G1 cyclins Cln1 and Cln2. SBF is kept inactive by co-repressor Whi5 until Cln3/CDK phosphorylates Whi5 to sequester the protein to the cytoplasm. As SBF is able to promote transcription, it activates *CLN1* and *CLN2* gene expression. Cln1/CDK and Cln2/CDK further inactivate

Whi5, leading to increased expression of *CLN1* and *CLN2* [159]. SBF activity is then inhibited by Clb2 later in the cell cycle [98]. The second positive feedback loop is between SFF (Swi Five Factor) and mitotic cyclin Clb2 [98]. Un-modified SFF has very little activity and activates transcription poorly if not activated. SFF activates expression of *CLB2* that when translated and bound to CDK, phosphorylates and further activates components of SFF [101, 102, 127]. One hypothesis drawn from studies eliminating cyclin/CDK activity states that the period of transcriptional oscillations increases as positive feedback loops are removed from the cells [133]. In cells with persistent Clb2/CDK activity, the SFF positive feedback loop is maintained while the SBF positive feedback loop is inhibited. This is the opposite case in cells lacking S-phase and mitotic cyclins, which maintain the SBF positive feedback while losing the SFF positive feedback. Perhaps both positive feedback loops are able to maintain transcriptional oscillations at the same period. This may explain why the period is the same in these two conditions, and the period is extended even further when both of these positive feedback loops are eliminated [133].

This novel experimental design has led to a deeper understanding of how a TF network functions to maintain transcriptional oscillations and of the relationship between cyclin/CDK activity and a TF network oscillator.

4.1.1 A new perspective on TF network structure and oscillations

The data presented in this dissertation challenges many of the preconceptions concerning the relationship between cyclin/CDK activity and a TF network oscillator. For many years, it has been proposed that a TF network could support oscillations, but its activity is regulated primarily through cyclin/CDK-mediated phosphorylation (reviewed in [5, 88, 90, 97]). Given the amount of cyclin/CDK feedback, it could be expected that when cell-cycle progression is arrested with persistent Clb2/CDK activity, transcriptional oscillations should also be arrested. After all, Clb2/CDK activity inhibits SBF, Ace2, and Swi5 (Figure 2.1 and 2.5) [98, 103, 104]. This should be sufficient to arrest transcription at G1/S and at M/G1. These are critical transitions for cells, as one represents commitment to the cell cycle and the other represents the transition from the end of one cell cycle into another cell cycle, conditions permitting. How could a transcriptional signal be passed along that could maintain oscillations?

While SBF gene expression is inhibited in these cells, MBF gene expression is largely unaffected (Figure 2.1, 2.2). This bifurcation in transcriptional activation at Start may be sufficient to trigger and maintain periodic transcription (Figure 1.3). Persistent Clb2/CDK activity also inhibits Ace2 and Swi5 activity via cytoplasmic sequestration [103, 104]. How is a signal passed from the end of one transcriptional cycle into the beginning of the next cycle in the absence of Ace2 and Swi5? Previous studies have shown that cells

are viable in the absence of both Ace2 and Swi5 [160], suggesting the presence of some yet unknown TF that plays a role in controlling periodic transcription at this transition. There is still much work left to understand the elements of the TF network that plays a role in propagating oscillations and periodic transcription. However, this experimental condition reveals novel ways to think about how a TF network functions to maintain periodic transcription.

4.2 A new mechanism of regulating a TF network oscillator

Previous work and studies completed in this dissertation have shown that regardless of how cell-cycle progression and oscillations in cyclin/CDK activity are arrested, transcriptional oscillations are maintained (Chapter 2) [81, 133]. The model proposing that a TF network acts as the cell-cycle oscillator that controls the timing of events fits with this data. However, if a TF network is in fact responsible for regulating cell-cycle events, shouldn't the TF network be tied to cell-cycle progression?

In all of the described experimental conditions, cell-cycle progression is halted while periodic transcription continues, suggesting that the TF network is maintaining oscillations. All of these experimental conditions used mutants to eliminate oscillations in cyclin/CDK activity, and as a result arrests cell-cycle events as well. In wild-type cells, cell-cycle events can be disrupted by environmental insults and intracellular perturbations. To maintain proper temporal order of events, signaling pathways called checkpoints monitor progress of

events and arrest later events if an earlier event is unable to be completed. How does the TF network and periodic transcription behave in response to checkpoints arresting cell-cycle progression and oscillations in cyclin/CDK activity?

When the DNA replication checkpoint or spindle assembly checkpoint is triggered, the bulk of periodic transcription is arrested (Figure 2.10). This finding suggests that the TF network is also arrested and leads to a new control module that is able to regulate the function of the TF network oscillator during the cell cycle (Figure 4.1).

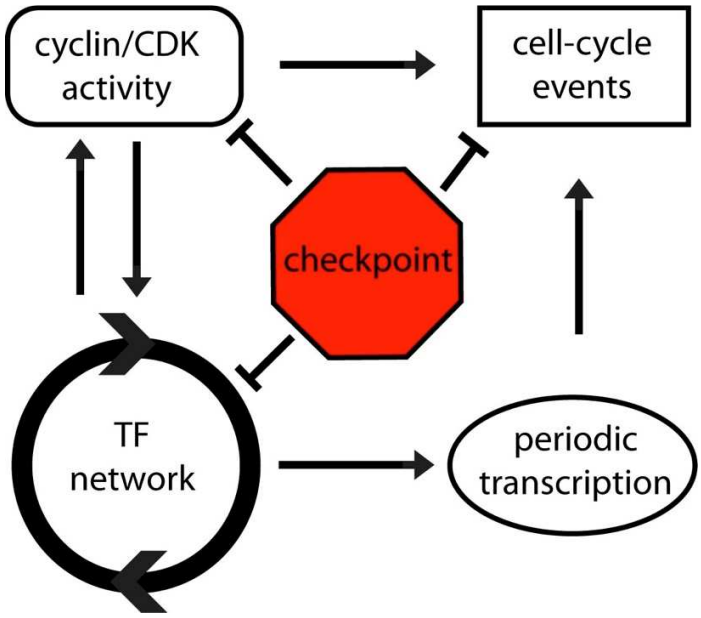

Figure 4.1: Checkpoints add another regulatory component to regulating a transcription factor network.

For the first time, cell-cycle progression, cyclin/CDK activity, and the TF network oscillator are coupled together. This result is remarkable as no study has demonstrated such a large transcriptional response to cell-cycle checkpoints. A number of genes have been found to be expressed during the DNA replication checkpoint [122], but it was proposed that this transcriptional response is meant to aid in overcoming replicative stress, not coupling periodic transcription to cell-cycle progression. Even more remarkably, no study has ever shown that the spindle assembly checkpoint triggers a transcriptional response. Given this striking transcriptional arrest, understanding the mechanism by which these two different checkpoints inhibit TF network oscillations my lead to understanding how the TF network functions during in normally cycling cells.

4.2.1 Potential checkpoint effectors that regulate TF network oscillations

The DNA replication checkpoint and spindle assembly checkpoint both arrest cell-cycle progression by inhibiting chromosome segregation. This action is prevented by inhibiting Cdc20, an activator of the E3 ubiquitin ligase APC [43, 60-63, 140]. Inhibiting APC^{Cdc20} activity stabilizes securin (Pds1 in budding yeast) by inhibiting separase (Esp1 in budding yeast) [140]. This results in maintained cohesion complexes around sister chromatids, preventing premature segregation. Another result of this mechanism is stabilizing other APC^{Cdc20} targets. The most obvious APC target that is affected is Clb2 (Figure 2.13).

Clb2/CDK can phosphorylate and activate APC, and with some delay, APC ubiquitinates and marks Clb2 for destruction [140]. This negative feedback loop serves as the basis for cell-cycle oscillations in early developing embryos. With Cdc20 inhibited during both of these checkpoints, Clb2 becomes stabilized and Clb2/CDK activity is persistent [47, 48]. This outcome is the same as the experiment I designed to test the effect of persistent mitotic cyclin/CDK activity on periodic transcription.

While only arresting cell-cycle progression by inhibiting APC^{Cdc20} results in continued transcriptional oscillations (Figure 2.2), arresting cells with checkpoints halts the bulk of periodic transcription (Figure 2.3). This result suggests that checkpoint-specific effectors play some role in arresting TF network oscillations and periodic transcription. What effectors could these be? And how do they impinge on the activity of network TFs?

The checkpoint kinases Rad53 and Dun1 are responsible for activating the transcriptional response during the DNA replication checkpoint (reviewed in [59]). While Dun1 has not been shown to affect any network TFs, it is possible that Dun1 not only triggers a transcriptional response unique to the DNA replication checkpoint, but also affects periodic transcription. Further, Rad53 has been shown to phosphorylate and alter the activity of several network TFs. Rad53 has been shown to phosphorylate and inactivate Swi6, a component of both SBF and MBF [96, 121]. Additionally, Rad53 inhibits Nrm1, a target and co-

repressor of MBF [122]. It is possible that by inhibiting transcriptional oscillations at MBF via Nrm1 and SBF via Swi6 and Clb2/CDK is sufficient to arrest downstream periodic transcription. Further experiments that perturb individual checkpoint kinases while a checkpoint is triggered will help to elucidate the contribution of each effector.

As mentioned above, no transcriptional changes in response to the spindle assembly checkpoint has yet to be reported outside of the work presented in this dissertation. Outside of the transcriptional changes observed due to Clb2 stabilization, it is unknown what checkpoint effectors in this signaling pathway could play a role in regulating the TF network oscillator. There are a number of candidates, including the many kinases that may affect network TFs. Using a candidate approach and an unbiased genetic screen may be one method to identify potential regulators. Better understanding how both of these checkpoints regulate the activity of the TF network oscillator may lead to an even better understanding of the topology of the network and how it can be regulated during the cell cycle.

4.3 Clarifying transcriptional regulation at Start

The massive amount of data generated to address the question of how cell-cycle regulated transcription is controlled during the cell-cycle provides much more information than what this single question uses. These datasets are so rich with extra information about the biology of the cell cycle and how different groups

of genes are regulated under a variety of conditions. Prior to these genomics studies, it was proposed that a TF network could support periodic transcription using genome-wide binding information [107]. However, physical binding does not confirm a functional effect on gene expression. Perhaps one of the most well-studied clusters of genes are those activated at Start. Redundant TF complexes SBF and MBF regulate these G1/S genes. SBF and MBF are heterodimers composed of a DNA binding protein, Swi4 and Mbp1, respectively and a cofactor, Swi6 that is a component of both complexes. Each have very similar consensus binding sequences. Genetic studies demonstrate that cells are viable in *swi4* and *mbp1* single mutants, but inviable when both are deleted [110], suggesting that SBF and MBF are highly redundant with each other.

Several genome-wide binding studies have attempted to elucidate the overlapping and exclusive SBF and MBF targets [106, 112, 113, 161]. While all of these studies propose that each TF complex does have a set of unique targets, agreement from these studies is very low. Further, more targeted studies using ChIP-PCR have shown that the overlap of shared targets may be much more extensive than once believed [111]. Can a functional genomics approach utilizing the datasets already generated, aid in clarifying the targets of SBF and MBF?

The experimental conditions used to probe the function of the TF network and the periodic transcriptional response may provide an opportunity to delineate

SBF targets, MBF targets, and shared targets. This can be done due to the different mechanisms utilized to inactivate SBF and MBF activity during the cell cycle. As mentioned above, SBF is inactivated via Clb2/CDK [98], while MBF is inactivated via a negative feedback loop with its own target Nrm1 [119]. In the absence of S-phase and mitotic cyclins (CDK "off"; *clb1,2,3,4,5,6*), canonical SBF targets have elevated gene expression levels as Clb2/CDK cannot inhibit SBF activity. However, since MBF is not regulated by S-phase or mitotic cyclins, its targets' gene expression dynamics are not affected in these cells (Figure 4.2). Similarly, in cells with persistent mitotic cyclin/CDK activity (CDK "on"; P_{GALL}-*CDC20*), SBF targets are inhibited after one cycle of expression while canonical MBF target gene expression is unaffected for the same reasons as in CDK "off" cells (Figure 4.2). A third experimental condition in which SBF and MBF target transcript dynamics differ is during the DNA replication checkpoint. As described above, this checkpoint arrests cell-cycle progression with persistent Clb2/CDK activity [47, 48], resulting in only one cycle of gene expression for SBF targets (Figure 4.2). This checkpoint also inactivates Nrm1 via Rad53-dependent phosphorylation [122], leading to persistent MBF activity and MBF target gene expression (Figure 4.2). These differences in regulation may lead to a better understanding of how this cluster of genes is activated by two redundant TF complexes.

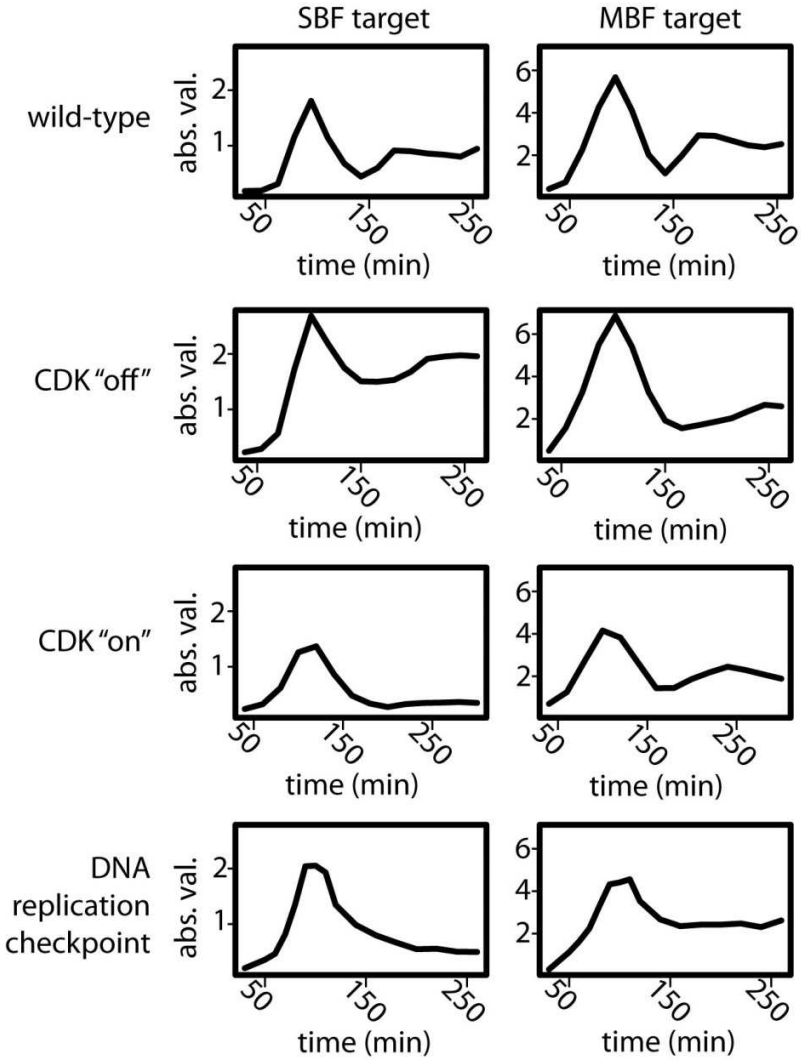

Figure 4.2: SBF and MBF targets are regulated differently during different experimental conditions. Absolute mRNA levels (dChip-normalized Affymetrix intensity units/1000) are shown for SBF target, *CLN1* (left) and MBF target, *POL1* (right) across different experimental conditions.

While this functional genomics approach seems very straightforward, complex combinatorial control could demonstrate that the transcriptional control at Start is much more complicated than anticipated. A recent study has shown

that a number of genes are activated by SBF, but inhibited by MBF [162]. Complex logic like this example may not be the only instance of how G1/S genes are controlled using combinatorial logics. Other factors might also play a role. For example, the number of binding sites or the order of the binding sites could affect which TF complex plays a dominant role in activating or repressing target gene expression. To tackle this interesting question, a computational approach should be taken to use all of the bioinformatics resources available to sort out groups of similarly controlled genes within the larger umbrella of SBF- and MBF-mediated gene expression.

4.4 Concluding remarks

In this dissertation, I have probed the mechanisms that play a role in regulating the TF network oscillator during the cell-cycle. I have shown that Clb2/CDK activity does in fact play a major role in regulating the transcriptional output of many network TFs. However, I have also shown that this regulation is not sufficient to arrest the TF network oscillator. These data are consistent with a new model of cell-cycle regulation centered on this TF network oscillator. I have further shown how periodic transcription is coupled to other cell-cycle events and control modules during a wild-type cell cycle. While there are many open questions that have come out of this work, it has contributed to a better understanding of how the cell cycle is regulated by multiple layers of control

modules and will hopefully contribute to new and exciting perspectives in the cell-cycle field.

References

1. Rao, P.N. and R.T. Johnson, *Mammalian cell fusion: studies on the regulation of DNA synthesis and mitosis.* Nature, 1970. **225**(5228): p. 159-64.

2. Draetta, G., et al., *Cdc2 protein kinase is complexed with both cyclin A and B: evidence for proteolytic inactivation of MPF.* Cell, 1989. **56**(5): p. 829-38.

3. Labbe, J.C., et al., *MPF from starfish oocytes at first meiotic metaphase is a heterodimer containing one molecule of cdc2 and one molecule of cyclin B.* EMBO J, 1989. **8**(10): p. 3053-8.

4. Morgan, D.O., *Cyclin-dependent kinases: engines, clocks, and microprocessors.* Annu Rev Cell Dev Biol, 1997. **13**: p. 261-91.

5. Morgan, D.O., *The cell cycle : principles of control*. Primers in biology. 2007, London
Sunderland, MA: Published by New Science Press in association with Oxford University Press ;
Distributed inside North America by Sinauer Associates, Publishers. xxvii, 297 p.

6. Evans, T., et al., *Cyclin: a protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division.* Cell, 1983. **33**(2): p. 389-96.

7. King, R.W., et al., *A 20S complex containing CDC27 and CDC16 catalyzes the mitosis-specific conjugation of ubiquitin to cyclin B.* Cell, 1995. **81**(2): p. 279-88.

8. Sudakin, V., et al., *The cyclosome, a large complex containing cyclin-selective ubiquitin ligase activity, targets cyclins for destruction at the end of mitosis.* Mol Biol Cell, 1995. **6**(2): p. 185-97.

9. Xiong, Y., et al., *Human D-type cyclin.* Cell, 1991. **65**(4): p. 691-9.

10. Swenson, K.I., K.M. Farrell, and J.V. Ruderman, *The clam embryo protein cyclin A induces entry into M phase and the resumption of meiosis in Xenopus oocytes.* Cell, 1986. **47**(6): p. 861-70.

11. Lew, D.J., V. Dulic, and S.I. Reed, *Isolation of three novel human cyclins by rescue of G1 cyclin (Cln) function in yeast.* Cell, 1991. **66**(6): p. 1197-206.

12. Culotti, J. and L.H. Hartwell, *Genetic control of the cell division cycle in yeast. 3. Seven genes controlling nuclear division.* Exp Cell Res, 1971. **67**(2): p. 389-401.

13. Hartwell, L.H., *Genetic control of the cell division cycle in yeast. IV. Genes controlling bud emergence and cytokinesis.* Exp Cell Res, 1971. **69**(2): p. 265-76.

14. Hartwell, L.H., *Genetic control of the cell division cycle in yeast. II. Genes controlling DNA replication and its initiation.* J Mol Biol, 1971. **59**(1): p. 183-94.

15. Hartwell, L.H., J. Culotti, and B. Reid, *Genetic control of the cell-division cycle in yeast. I. Detection of mutants.* Proc Natl Acad Sci U S A, 1970. **66**(2): p. 352-9.

16. Hartwell, L.H., et al., *Genetic Control of the Cell Division Cycle in Yeast: V. Genetic Analysis of cdc Mutants.* Genetics, 1973. **74**(2): p. 267-86.

17. Nurse, P., P. Thuriaux, and K. Nasmyth, *Genetic control of the cell division cycle in the fission yeast Schizosaccharomyces pombe.* Mol Gen Genet, 1976. **146**(2): p. 167-78.

18. Reed, S.I., J.A. Hadwiger, and A.T. Lorincz, *Protein kinase activity associated with the product of the yeast cell division cycle gene CDC28.* Proc Natl Acad Sci U S A, 1985. **82**(12): p. 4055-9.

19. Simanis, V. and P. Nurse, *The cell cycle control gene cdc2+ of fission yeast encodes a protein kinase potentially regulated by phosphorylation.* Cell, 1986. **45**(2): p. 261-8.

20. Hindley, J. and G.A. Phear, *Sequence of the cell division gene CDC2 from Schizosaccharomyces pombe; patterns of splicing and homology to protein kinases.* Gene, 1984. **31**(1-3): p. 129-34.

21. Lorincz, A.T. and S.I. Reed, *Primary structure homology between the product of yeast cell division control gene CDC28 and vertebrate oncogenes.* Nature, 1984. **307**(5947): p. 183-5.

22. Beach, D., M. Piper, and P. Nurse, *Construction of a Schizosaccharomyces pombe gene bank in a yeast bacterial shuttle vector and its use to isolate genes by complementation.* Mol Gen Genet, 1982. **187**(2): p. 326-9.

23. Lee, M.G. and P. Nurse, *Complementation used to clone a human homologue of the fission yeast cell cycle control gene cdc2.* Nature, 1987. **327**(6117): p. 31-5.

24. Lamb, J.R., et al., *Cdc16p, Cdc23p and Cdc27p form a complex essential for mitosis.* EMBO J, 1994. **13**(18): p. 4321-8.

25. Zachariae, W., et al., *Identification of subunits of the anaphase-promoting complex of Saccharomyces cerevisiae.* Science, 1996. **274**(5290): p. 1201-4.

26. Murray, A.W. and M.W. Kirschner, *Dominoes and clocks: the union of two views of the cell cycle.* Science, 1989. **246**(4930): p. 614-21.

27. Dahmann, C., J.F. Diffley, and K.A. Nasmyth, *S-phase-promoting cyclin-dependent kinases prevent re-replication by inhibiting the transition of replication origins to a pre-replicative state.* Curr Biol, 1995. **5**(11): p. 1257-69.

28. Haase, S.B., M. Winey, and S.I. Reed, *Multi-step control of spindle pole body duplication by cyclin-dependent kinase.* Nat Cell Biol, 2001. **3**(1): p. 38-42.

29. Simmons Kovacs, L.A., C.L. Nelson, and S.B. Haase, *Intrinsic and cyclin-dependent kinase-dependent control of spindle pole body duplication in budding yeast.* Mol Biol Cell, 2008. **19**(8): p. 3243-53.

30. Hartwell, L.H. and T.A. Weinert, *Checkpoints: controls that ensure the order of cell cycle events.* Science, 1989. **246**(4930): p. 629-34.

31. Lew, D.J., *The morphogenesis checkpoint: how yeast cells watch their figures.* Curr Opin Cell Biol, 2003. **15**(6): p. 648-53.

32. Elledge, S.J., *Cell cycle checkpoints: preventing an identity crisis.* Science, 1996. **274**(5293): p. 1664-72.

33. Clarke, D.J. and J.F. Gimenez-Abian, *Checkpoints controlling mitosis.* Bioessays, 2000. **22**(4): p. 351-63.

34. Lew, D.J. and D.J. Burke, *The spindle assembly and spindle position checkpoints.* Annu Rev Genet, 2003. **37**: p. 251-82.

35. Lopes, M., et al., *The DNA replication checkpoint response stabilizes stalled replication forks.* Nature, 2001. **412**(6846): p. 557-61.

36. Tourriere, H. and P. Pasero, *Maintenance of fork integrity at damaged DNA and natural pause sites.* DNA Repair (Amst), 2007. **6**(7): p. 900-13.

37. Araki, H., et al., *Dpb11, which interacts with DNA polymerase II(epsilon) in Saccharomyces cerevisiae, has a dual role in S-phase progression and at a cell cycle checkpoint.* Proc Natl Acad Sci U S A, 1995. **92**(25): p. 11791-5.

38. Navas, T.A., Z. Zhou, and S.J. Elledge, *DNA polymerase epsilon links the DNA replication machinery to the S phase checkpoint.* Cell, 1995. **80**(1): p. 29-39.

39. Sugimoto, K., et al., *Rfc5, a small subunit of replication factor C complex, couples DNA replication and mitosis in budding yeast.* Proc Natl Acad Sci U S A, 1996. **93**(14): p. 7048-52.

40. Wang, H. and S.J. Elledge, *DRC1, DNA replication and checkpoint protein 1, functions with DPB11 to control DNA replication and the S-phase checkpoint in Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 1999. **96**(7): p. 3824-9.

41. Weinert, T.A., G.L. Kiser, and L.H. Hartwell, *Mitotic checkpoint genes in budding yeast and the dependence of mitosis on DNA replication and repair.* Genes Dev, 1994. **8**(6): p. 652-65.

42. Kato, R. and H. Ogawa, *An essential gene, ESR1, is required for mitotic cell growth, DNA repair and meiotic recombination in Saccharomyces cerevisiae.* Nucleic Acids Res, 1994. **22**(15): p. 3104-12.

43. Osborn, A.J. and S.J. Elledge, *Mrc1 is a replication fork component whose phosphorylation in response to DNA replication stress activates Rad53.* Genes Dev, 2003. **17**(14): p. 1755-67.

44. Agarwal, R., et al., *Two distinct pathways for inhibiting pds1 ubiquitination in response to DNA damage.* J Biol Chem, 2003. **278**(45): p. 45027-33.

45. Yamamoto, A., V. Guacci, and D. Koshland, *Pds1p, an inhibitor of anaphase in budding yeast, plays a critical role in the APC and checkpoint pathway(s).* J Cell Biol, 1996. **133**(1): p. 99-110.

46. Cohen-Fix, O., et al., *Anaphase initiation in Saccharomyces cerevisiae is controlled by the APC-dependent degradation of the anaphase inhibitor Pds1p.* Genes Dev, 1996. **10**(24): p. 3081-93.

47. Sorger, P.K. and A.W. Murray, *S-phase feedback control in budding yeast independent of tyrosine phosphorylation of p34cdc28.* Nature, 1992. **355**(6358): p. 365-8.

48. Stueland, C.S., et al., *Full activation of p34CDC28 histone H1 kinase activity is unable to promote entry into mitosis in checkpoint-arrested cells of the yeast Saccharomyces cerevisiae.* Mol Cell Biol, 1993. **13**(6): p. 3744-55.

49. Surana, U., et al., *Destruction of the CDC28/CLB mitotic kinase is not required for the metaphase to anaphase transition in budding yeast.* EMBO J, 1993. **12**(5): p. 1969-78.

50. Bashkirov, V.I., et al., *Direct kinase-to-kinase signaling mediated by the FHA phosphoprotein recognition domain of the Dun1 DNA damage checkpoint kinase.* Mol Cell Biol, 2003. **23**(4): p. 1441-52.

51. Zhao, X. and R. Rothstein, *The Dun1 checkpoint kinase phosphorylates and regulates the ribonucleotide reductase inhibitor Sml1.* Proc Natl Acad Sci U S A, 2002. **99**(6): p. 3746-51.

52. Chabes, A., V. Domkin, and L. Thelander, *Yeast Sml1, a protein inhibitor of ribonucleotide reductase.* J Biol Chem, 1999. **274**(51): p. 36679-83.

53. Zhou, Z. and S.J. Elledge, *DUN1 encodes a protein kinase that controls the DNA damage response in yeast.* Cell, 1993. **75**(6): p. 1119-27.

54. Huang, M., Z. Zhou, and S.J. Elledge, *The DNA replication and damage checkpoint pathways induce transcription by inhibition of the Crt1 repressor.* Cell, 1998. **94**(5): p. 595-605.

55. Keleher, C.A., et al., *Ssn6-Tup1 is a general repressor of transcription in yeast.* Cell, 1992. **68**(4): p. 709-19.

56. Zhou, Z. and S.J. Elledge, *Isolation of crt mutants constitutive for transcription of the DNA damage inducible gene RNR3 in Saccharomyces cerevisiae.* Genetics, 1992. **131**(4): p. 851-66.

57. McClanahan, T. and K. McEntee, *Specific transcripts are elevated in Saccharomyces cerevisiae in response to DNA damage.* Mol Cell Biol, 1984. **4**(11): p. 2356-63.

58. Ruby, S.W. and J.W. Szostak, *Specific Saccharomyces cerevisiae genes are expressed in response to DNA-damaging agents.* Mol Cell Biol, 1985. **5**(1): p. 75-84.

59. Fu, Y., L. Pastushok, and W. Xiao, *DNA damage-induced gene expression in Saccharomyces cerevisiae.* FEMS Microbiol Rev, 2008. **32**(6): p. 908-26.

60. Fang, G., H. Yu, and M.W. Kirschner, *The checkpoint protein MAD2 and the mitotic regulator CDC20 form a ternary complex with the anaphase-promoting complex to control anaphase initiation.* Genes Dev, 1998. **12**(12): p. 1871-83.

61. Hwang, L.H., et al., *Budding yeast Cdc20: a target of the spindle checkpoint.* Science, 1998. **279**(5353): p. 1041-4.

62. Luo, X., et al., *Structure of the Mad2 spindle assembly checkpoint protein and its interaction with Cdc20.* Nat Struct Biol, 2000. **7**(3): p. 224-9.

63. Hardwick, K.G., et al., *MAD3 encodes a novel component of the spindle checkpoint which interacts with Bub3p, Cdc20p, and Mad2p.* J Cell Biol, 2000. **148**(5): p. 871-82.

64. Hereford, L.M., et al., *Cell-cycle regulation of yeast histone mRNA.* Cell, 1981. **24**(2): p. 367-75.

65. Nasmyth, K., *Molecular analysis of a cell lineage.* Nature, 1983. **302**(5910): p. 670-6.

66. Storms, R.K., et al., *Cell cycle-dependent expression of thymidylate synthase in Saccharomyces cerevisiae.* Mol Cell Biol, 1984. **4**(12): p. 2858-64.

67. White, J.H.M., et al., *Periodic Transcription as a Means of Regulating Gene-Expression during the Cell-Cycle - Contrasting Modes of Expression of DNA-Ligase Genes in Budding and Fission Yeast.* Embo Journal, 1986. **5**(7): p. 1705-1709.

68. Kupiec, M. and G. Simchen, *Regulation of the RAD6 gene of Saccharomyces cerevisiae in the mitotic cell cycle and in meiosis.* Mol Gen Genet, 1986. **203**(3): p. 538-43.

69. Nasmyth, K., A. Seddon, and G. Ammerer, *Cell cycle regulation of SW15 is required for mother-cell-specific HO transcription in yeast.* Cell, 1987. **49**(4): p. 549-58.

70. White, J.H.M., et al., *The Cdc8 Transcript Is Cell-Cycle Regulated in Yeast and Is Expressed Coordinately with Cdc9 and Cdc21 at a Point Preceding Histone Transcription.* Experimental Cell Research, 1987. **171**(1): p. 223-231.

71. Johnston, L.H., et al., *The yeast DNA polymerase I transcript is regulated in both the mitotic cell cycle and in meiosis and is also induced after DNA damage.* Nucleic Acids Res, 1987. **15**(13): p. 5017-30.

72. Chapman, J.W. and L.H. Johnston, *The yeast gene, DBF4, essential for entry into S phase is cell cycle regulated.* Exp Cell Res, 1989. **180**(2): p. 419-28.

73. Johnston, L.H., et al., *Expression of the yeast DNA primase gene, PRI1, is regulated within the mitotic cell cycle and in meiosis.* Mol Gen Genet, 1990. **221**(1): p. 44-8.

74. Johnston, L.H., et al., *The product of the Saccharomyces cerevisiae cell cycle gene DBF2 has homology with protein kinases and is periodically expressed in the cell cycle.* Mol Cell Biol, 1990. **10**(4): p. 1358-66.

75. Price, C., K. Nasmyth, and T. Schuster, *A general approach to the isolation of cell cycle-regulated genes in the budding yeast, Saccharomyces cerevisiae.* J Mol Biol, 1991. **218**(3): p. 543-56.

76. Ramsay, G., *DNA chips: state-of-the art.* Nat Biotechnol, 1998. **16**(1): p. 40-4.

77. Cho, R.J., et al., *A genome-wide transcriptional analysis of the mitotic cell cycle.* Mol Cell, 1998. **2**(1): p. 65-73.

78. Spellman, P.T., et al., *Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization.* Mol Biol Cell, 1998. **9**(12): p. 3273-97.

79. de Lichtenberg, U., et al., *Comparison of computational methods for the identification of cell cycle-regulated genes.* Bioinformatics, 2005. **21**(7): p. 1164-71.

80. Pramila, T., et al., *The Forkhead transcription factor Hcm1 regulates chromosome segregation genes and fills the S-phase gap in the transcriptional circuitry of the cell cycle.* Genes Dev, 2006. **20**(16): p. 2266-78.

81. Orlando, D.A., et al., *Global control of cell-cycle transcription by coupled CDK and network oscillators.* Nature, 2008. **453**(7197): p. 944-7.

82. Rustici, G., et al., *Periodic gene expression program of the fission yeast cell cycle.* Nat Genet, 2004. **36**(8): p. 809-17.

83. Peng, X., et al., *Identification of cell cycle-regulated genes in fission yeast.* Mol Biol Cell, 2005. **16**(3): p. 1026-42.

84. Oliva, A., et al., *The cell cycle-regulated genes of Schizosaccharomyces pombe.* PLoS Biol, 2005. **3**(7): p. e225.

85. Cho, R.J., et al., *Transcriptional regulation and function during the human cell cycle.* Nat Genet, 2001. **27**(1): p. 48-54.

86. Whitfield, M.L., et al., *Identification of genes periodically expressed in the human cell cycle and their expression in tumors.* Mol Biol Cell, 2002. **13**(6): p. 1977-2000.

87. Futcher, B., *Microarrays and cell cycle transcription in yeast.* Curr Opin Cell Biol, 2000. **12**(6): p. 710-5.

88. McInerny, C.J., *Cell cycle regulated gene expression in yeasts.* Adv Genet, 2011. **73**: p. 51-85.

89. Simmons Kovacs, L.A., D.A. Orlando, and S.B. Haase, *Transcription networks and cyclin/CDKs: the yin and yang of cell cycle oscillators.* Cell Cycle, 2008. **7**(17): p. 2626-9.

90. Breeden, L.L., *Periodic transcription: a cycle within a cycle.* Curr Biol, 2003. **13**(1): p. R31-8.

91. Verma, R., et al., *Phosphorylation of Sic1p by G1 Cdk required for its degradation and entry into S phase.* Science, 1997. **278**(5337): p. 455-60.

92. Knapp, D., et al., *The transcription factor Swi5 regulates expression of the cyclin kinase inhibitor p40SIC1.* Mol Cell Biol, 1996. **16**(10): p. 5701-7.

93. Schwob, E., et al., *The B-type cyclin kinase inhibitor p40SIC1 controls the G1 to S transition in S. cerevisiae.* Cell, 1994. **79**(2): p. 233-44.

94. Feldman, R.M., et al., *A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p.* Cell, 1997. **91**(2): p. 221-30.

95. Barberis, M., T.W. Spiesser, and E. Klipp, *Replication origins and timing of temporal replication in budding yeast: how to solve the conundrum?* Curr Genomics, 2010. **11**(3): p. 199-211.

96. Sidorova, J.M., G.E. Mikesell, and L.L. Breeden, *Cell cycle-regulated phosphorylation of Swi6 controls its nuclear localization.* Mol Biol Cell, 1995. **6**(12): p. 1641-58.

97. Wittenberg, C. and S.I. Reed, *Cell cycle-dependent transcription in yeast: promoters, transcription factors, and transcriptomes.* Oncogene, 2005. **24**(17): p. 2746-55.

98. Amon, A., et al., *Mechanisms that help the yeast cell cycle clock tick: G2 cyclins transcriptionally activate G2 cyclins and repress G1 cyclins.* Cell, 1993. **74**(6): p. 993-1007.

99. Ubersax, J.A., et al., *Targets of the cyclin-dependent kinase Cdk1.* Nature, 2003. **425**(6960): p. 859-64.

100. Ostapenko, D., J.L. Burton, and M.J. Solomon, *Identification of anaphase promoting complex substrates in S. cerevisiae.* PLoS One, 2012. **7**(9): p. e45895.

101. Darieva, Z., et al., *Cell cycle-regulated transcription through the FHA domain of Fkh2p and the coactivator Ndd1p.* Curr Biol, 2003. **13**(19): p. 1740-5.

102. Reynolds, D., et al., *Recruitment of Thr 319-phosphorylated Ndd1p to the FHA domain of Fkh2p requires Clb kinase activity: a mechanism for CLB cluster gene activation.* Genes Dev, 2003. **17**(14): p. 1789-802.

103. O'Conallain, C., et al., *Regulated nuclear localisation of the yeast transcription factor Ace2p controls expression of chitinase (CTS1) in Saccharomyces cerevisiae.* Mol Gen Genet, 1999. **262**(2): p. 275-82.

104. Moll, T., et al., *The role of phosphorylation and the CDC28 protein kinase in cell cycle-regulated nuclear import of the S. cerevisiae transcription factor SWI5.* Cell, 1991. **66**(4): p. 743-58.

105. Harbison, C.T., et al., *Transcriptional regulatory code of a eukaryotic genome.* Nature, 2004. **431**(7004): p. 99-104.

106. Lee, T.I., et al., *Transcriptional regulatory networks in Saccharomyces cerevisiae.* Science, 2002. **298**(5594): p. 799-804.

107. Simon, I., et al., *Serial regulation of transcriptional regulators in the yeast cell cycle.* Cell, 2001. **106**(6): p. 697-708.

108. Dirick, L., et al., *A central role for SWI6 in modulating cell cycle Start-specific transcription in yeast.* Nature, 1992. **357**(6378): p. 508-13.

109. Andrews, B.J. and I. Herskowitz, *The yeast SWI4 protein contains a motif present in developmental regulators and is part of a complex involved in cell-cycle-dependent transcription.* Nature, 1989. **342**(6251): p. 830-3.

110. Koch, C., et al., *A role for the transcription factors Mbp1 and Swi4 in progression from G1 to S phase.* Science, 1993. **261**(5128): p. 1551-7.

111. Bean, J.M., E.D. Siggia, and F.R. Cross, *High functional overlap between MluI cell-cycle box binding factor and Swi4/6 cell-cycle box binding factor in the G1/S transcriptional program in Saccharomyces cerevisiae.* Genetics, 2005. **171**(1): p. 49-61.

112. Ferrezuelo, F., et al., *The transcriptional network activated by Cln3 cyclin at the G1-to-S transition of the yeast cell cycle.* Genome Biol, 2010. **11**(6): p. R67.

113. Iyer, V.R., et al., *Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF.* Nature, 2001. **409**(6819): p. 533-8.

114. Costanzo, M., et al., *CDK activity antagonizes Whi5, an inhibitor of G1/S transcription in yeast.* Cell, 2004. **117**(7): p. 899-913.

115. Cross, F.R. and A.H. Tinkelenberg, *A potential positive feedback loop controlling CLN1 and CLN2 gene expression at the start of the yeast cell cycle.* Cell, 1991. **65**(5): p. 875-83.

116. de Bruin, R.A., et al., *Cln3 activates G1-specific transcription via phosphorylation of the SBF bound repressor Whi5.* Cell, 2004. **117**(7): p. 887-98.

117. Dirick, L. and K. Nasmyth, *Positive feedback in the activation of G1 cyclins in yeast.* Nature, 1991. **351**(6329): p. 754-7.

118. Pramila, T., et al., *Conserved homeodomain proteins interact with MADS box protein Mcm1 to restrict ECB-dependent transcription to the M/G1 phase of the cell cycle.* Genes Dev, 2002. **16**(23): p. 3034-45.

119. de Bruin, R.A., et al., *Constraining G1-specific transcription to late G1 phase: the MBF-associated corepressor Nrm1 acts via negative feedback.* Mol Cell, 2006. **23**(4): p. 483-96.

120. Sidorova, J.M. and L.L. Breeden, *Rad53-dependent phosphorylation of Swi6 and down-regulation of CLN1 and CLN2 transcription occur in response to DNA damage in Saccharomyces cerevisiae.* Genes Dev, 1997. **11**(22): p. 3032-45.

121. Sidorova, J.M. and L.L. Breeden, *Rad53 checkpoint kinase phosphorylation site preference identified in the Swi6 protein of Saccharomyces cerevisiae.* Mol Cell Biol, 2003. **23**(10): p. 3405-16.

122. Travesa, A., et al., *DNA replication stress differentially regulates G1/S genes via Rad53-dependent inactivation of Nrm1.* EMBO J, 2012. **31**(7): p. 1811-22.

123. Kumar, R., et al., *Forkhead transcription factors, Fkh1p and Fkh2p, collaborate with Mcm1p to control transcription required for M-phase.* Curr Biol, 2000. **10**(15): p. 896-906.

124. Pic, A., et al., *The forkhead protein Fkh2 is a component of the yeast cell cycle transcription factor SFF.* EMBO J, 2000. **19**(14): p. 3750-61.

125. Maher, M., et al., *Cell cycle-regulated transcription of the CLB2 gene is dependent on Mcm1 and a ternary complex factor.* Mol Cell Biol, 1995. **15**(6): p. 3129-37.

126. Lim, F.L., et al., *Mcm1p-induced DNA bending regulates the formation of ternary transcription factor complexes.* Mol Cell Biol, 2003. **23**(2): p. 450-61.

127. Koranda, M., et al., *Forkhead-like transcription factors recruit Ndd1 to the chromatin of G2/M-specific promoters.* Nature, 2000. **406**(6791): p. 94-8.

128. Althoefer, H., et al., *Mcm1 is required to coordinate G2-specific transcription in Saccharomyces cerevisiae.* Mol Cell Biol, 1995. **15**(11): p. 5917-28.

129. Pic-Taylor, A., et al., *Regulation of cell cycle-specific gene expression through cyclin-dependent kinase-mediated phosphorylation of the*

forkhead transcription factor Fkh2p. Mol Cell Biol, 2004. **24**(22): p. 10036-46.

130. Di Talia, S., et al., *Daughter-specific transcription factors regulate cell size control in budding yeast.* PLoS Biol, 2009. **7**(10): p. e1000221.

131. Haase, S.B. and S.I. Reed, *Evidence that a free-running oscillator drives G1 events in the budding yeast cell cycle.* Nature, 1999. **401**(6751): p. 394-7.

132. Lu, Y. and F.R. Cross, *Periodic cyclin-Cdk activity entrains an autonomous Cdc14 release oscillator.* Cell, 2010. **141**(2): p. 268-79.

133. Simmons Kovacs, L.A., et al., *Cyclin-dependent kinases are regulators and effectors of oscillations driven by a transcription factor network.* Mol Cell, 2012. **45**(5): p. 669-79.

134. Oikonomou, C. and F.R. Cross, *Frequency control of cell cycle oscillators.* Curr Opin Genet Dev, 2010. **20**(6): p. 605-12.

135. Glass, L., *Synchronization and rhythmic processes in physiology.* Nature, 2001. **410**(6825): p. 277-84.

136. Winfree, A.T., *Biological rhythms and the behavior of populations of coupled oscillators.* J Theor Biol, 1967. **16**(1): p. 15-42.

137. Charvin, G., F.R. Cross, and E.D. Siggia, *Forced periodic expression of G1 cyclins phase-locks the budding yeast cell cycle.* Proc Natl Acad Sci U S A, 2009. **106**(16): p. 6632-7.

138. Murray, A.W., M.J. Solomon, and M.W. Kirschner, *The role of cyclin synthesis and degradation in the control of maturation promoting factor activity.* Nature, 1989. **339**(6222): p. 280-6.

139. Murray, A.W. and M.W. Kirschner, *Cyclin synthesis drives the early embryonic cell cycle.* Nature, 1989. **339**(6222): p. 275-80.

140. Shirayama, M., et al., *APC(Cdc20) promotes exit from mitosis by destroying the anaphase inhibitor Pds1 and cyclin Clb5.* Nature, 1999. **402**(6758): p. 203-7.

141. Wasch, R. and F.R. Cross, *APC-dependent proteolysis of the mitotic cyclin Clb2 is essential for mitotic exit.* Nature, 2002. **418**(6897): p. 556-62.

142. Glynn, E.F., J. Chen, and A.R. Mushegian, *Detecting periodic patterns in unevenly spaced gene expression time series using Lomb-Scargle periodograms.* Bioinformatics, 2006. **22**(3): p. 310-6.

143. Lomb, N.R., *Least-Squares Frequency-Analysis of Unequally Spaced Data.* Astrophysics and Space Science, 1976. **39**(2): p. 447-462.

144. Scargle, J.D., *Studies in astronomical time series analysis. II-Statistical aspects of spectral analysis of unevenly spaced data.* Astrophysical Journal, 1982. **263**: p. 835-853.

145. Collins, K.A., et al., *The overexpression of a Saccharomyces cerevisiae centromeric histone H3 variant mutant protein leads to a defect in kinetochore biorientation.* Genetics, 2007. **175**(2): p. 513-25.

146. Dutta, C., et al., *The DNA replication checkpoint directly regulates MBF-dependent G1/S transcription.* Mol Cell Biol, 2008. **28**(19): p. 5977-85.

147. de Bruin, R.A., et al., *DNA replication checkpoint promotes G1-S transcription by inactivating the MBF repressor Nrm1.* Proc Natl Acad Sci U S A, 2008. **105**(32): p. 11230-5.

148. Storey, J.D., *A direct approach to false discovery rates.* Journal of the Royal Statistical Society Series B-Statistical Methodology, 2002. **64**: p. 479-498.

149. Teixeira, M.C., et al., *The YEASTRACT database: a tool for the analysis of transcription regulatory associations in Saccharomyces cerevisiae.* Nucleic Acids Res, 2006. **34**(Database issue): p. D446-51.

150. Frey, B.J. and D. Dueck, *Clustering by passing messages between data points.* Science, 2007. **315**(5814): p. 972-6.

151. Lim, W.K., et al., *Comparative analysis of microarray normalization procedures: effects on reverse engineering gene networks.* Bioinformatics, 2007. **23**(13): p. i282-8.

152. Li, C. and W. Hung Wong, *Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application.* Genome Biol, 2001. **2**(8): p. RESEARCH0032.

153. Li, C. and W.H. Wong, *Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection.* Proc Natl Acad Sci U S A, 2001. **98**(1): p. 31-6.

154. Do, J.H. and D.K. Choi, *Normalization of microarray data: single-labeled and dual-labeled arrays.* Mol Cells, 2006. **22**(3): p. 254-61.

155. Gautier, L., et al., *affy--analysis of Affymetrix GeneChip data at the probe level.* Bioinformatics, 2004. **20**(3): p. 307-15.

156. Hughes, M.E., J.B. Hogenesch, and K. Kornacker, *JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets.* J Biol Rhythms, 2010. **25**(5): p. 372-80.

157. Gentleman, R.C., et al., *Bioconductor: open software development for computational biology and bioinformatics.* Genome Biol, 2004. **5**(10): p. R80.

158. Orlando, D.A., et al., *A probabilistic model for cell cycle distributions in synchrony experiments.* Cell Cycle, 2007. **6**(4): p. 478-88.

159. Skotheim, J.M., et al., *Positive feedback of G1 cyclins ensures coherent cell cycle entry.* Nature, 2008. **454**(7202): p. 291-6.

160. Dohrmann, P.R., et al., *Parallel pathways of gene regulation: homologous regulators SWI5 and ACE2 differentially control transcription of HO and chitinase.* Genes Dev, 1992. **6**(1): p. 93-104.

161. Simon, I., et al., *Serial Regulation of Transcriptional Regulators in the Yeast Cell Cycle.* Cell, 2001. **106**(6): p. 697-708.

162. Bastos de Oliveira, F.M., et al., *Linking DNA replication checkpoint to MBF cell-cycle transcription reveals a distinct class of G1/S genes.* EMBO J, 2012. **31**(7): p. 1798-810.

Biography

Sara Lynn Bristow was born November 20, 1984 in Wilmington, Delaware. She was raised in Wilmington, Delaware through high school. Sara attended Georgetown University in Washington, District of Colombia. While there, she pursued a degree in biology and a minor in theology. Sara learned invaluable techniques and critical thinking skills while working in the laboratory of Dr. William Hahn at Georgetown University and in the laboratory of Dr. Patrick Gaffney at University of Delaware Marine Center in Lewes, Delaware. She also explored research in science education, focusing her junior and senior years at Georgetown writing a thesis on the the effects of learning styles on student knowledge. She received her Bachelor of Science degree in May 2007, graduating *magna cum laude* with High Honors in Biology.

In August 2007, Sara matriculated into the Cell and Molecular Biology training program at Duke University in Durham, North Carolina. She joined the laboratory of Dr. Steven Haase and the University Program in Genetics and Genomics for her dissertation research, which focuses on cell-cycle regulated transcription in *Saccharomyces cerevisiae*. During graduate school, she has written one manuscript that is to be reviewed and authored a review on the field of cell-cycle regulated transcription. After completing her Ph.D, Sara plans to focus her efforts on science education as a teacher.